



# Étude mathématique et numérique des méthodes de réduction dimensionnelle de type POD et PGD

Marwan Saleh

## ► To cite this version:

Marwan Saleh. Étude mathématique et numérique des méthodes de réduction dimensionnelle de type POD et PGD. Mécanique des fluides [physics.class-ph]. Université de La Rochelle, 2015. Français. NNT : 2015LAROS004 . tel-01278980

**HAL Id: tel-01278980**

**<https://theses.hal.science/tel-01278980>**

Submitted on 25 Feb 2016

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



UNIVERSITÉ DE LA ROCHELLE

## THÈSE

pour l'obtention du Grade de  
**DOCTEUR de l'Université de La Rochelle**  
École Doctorale: **SI-MMEA, ED 522**  
Laboratoire: **LaSIE**  
Mention: **Mécanique des fluides**  
présentée par:

**Marwan SALEH<sup>1</sup>**

---

### Étude mathématique et numérique des méthodes de réduction dimensionnelle de type POD et PGD

---

Soutenue le 07 Mai 2015  
devant la commission d'Examen

#### Jury

##### **Rapporteurs :**

Antonio Falco  
Alexei Lozinski

Professeur, Universidad CEU Cardenal Herrera (Spain)  
Professeur, Université de Franche-Comté, Besançon

##### **Examineurs :**

Marc Garbey  
Baudouin Denis de Senneville  
Nicolas Papadakis  
Aziz Hamdouni  
Abdallah El Hamidi

Professeur, Université de Houston (USA)  
CR (CNRS) - HDR, Université de Bordeaux I  
CR (CNRS), Université de Bordeaux I  
Professeur, Université de La Rochelle  
MCF - HDR, Université de La Rochelle

---

1. N'hésitez pas de m'envoyer vos remarques à: [marwan.saleh@univ-lr.fr](mailto:marwan.saleh@univ-lr.fr)



*À mes parents*



# Remerciements

D’abord, je commence par le remerciement de l’équipe du LaSIE représentée par son directeur **Karim Ait Mokhtar** pour m’avoir accueilli au sein de ce laboratoire.

Ensuite, je tiens à adresser mes plus chaleureux remerciements à mon directeur de thèse **Abdallah El Hamidi**, d’avoir accepté de diriger ma thèse, ainsi que pour sa famille **Amira, Nour et Riad**. Durant ces trois années, il m’a transmis la passion de la recherche en mathématiques et n’a eu de cesse de m’encourager et de me soutenir durant cette période. J’en profite pour lui exprimer ici ma plus profonde gratitude et je lui souhaite la bonne vie avec sa jolie famille.

Mon co-directeur de thèse **Marc Garbey**, je vous adresse mes sincères remerciements pour les riches visio GDTs avec les collègues internationaux. Je vous remercie aussi de m’avoir invité aux États-Unis malgré que je n’ai pas pu y aller à cause de la réponse tardive du consulat. Et j’aimerais vous rejoindre par la suite pour travailler avec vous sur vos sujets qui m’intéressent beaucoup.

Je remercie évidemment, mon professeur **Mustapha Jazar**, la personne qui m’inspire ainsi que mes collègues et je remercie encore tous les membres de l’association LaSER qu’il l’a fondé.

Je voudrais remercier mes rapporteurs **Antonio Falco** et **Alexei Lozinski** d’avoir acceptés de rapporter ma thèse et pour l’intérêt qu’ils ont porté à mon travail.

Je remercie également **Baudouin Denis de Senneville, Nicolas Papadakis**, d’avoir fait partie du jury de ma thèse.

Je suis très honoré de faire partie de l’axe AB sous la direction d’**Aziz Hamdouni**. J’y ai appris beaucoup de choses grâce à Aziz ainsi que mes autres collègues.

Je remercie également **Mokhtar Kirane** pour ses discussions riches que l’on a pu échanger durant ces trois années.

Durant ma thèse, j’ai eu la chance de rencontrer en France mon amour **Elise**, après elle, la vie était beaucoup plus belle. Nous avons eu un joli **Omar**, Je vous remercie spécialement **Oum Omar**, et **Omar**.

J’adresse mes meilleurs remerciements à tous mes amis, surtout : **Bilal, Hassan, Saadallah, Jounaid, Ahmad, Djamal et Abdelrazzak**.

Enfin, je n’oublie pas de remercier mes parents, mes frères, mes soeurs et ma belle famille pour leur soutien. Je remercie plus particulièrement, Oncle **Ahmad** et Mamie **Thérèse**.



# Table des matières

Remerciements . . . . .	v
Table des matières . . . . .	vii
<b>1 Introduction et présentation des travaux</b>	<b>1</b>
1.1 Intoduction à la Décomposition par Projection Orthogonale (POD) définie sur un espace de Hilbert complexe . . . . .	2
1.2 Espaces de Hilbert complexes . . . . .	4
1.3 Opérateurs de Hilbert-Schmidt . . . . .	6
1.4 Intégrales de Bochner . . . . .	8
1.5 Décomposition par projection orthogonale POD . . . . .	9
1.5.1 Méthode directe . . . . .	9
1.5.2 Méthode des snapshots . . . . .	12
1.6 Equivalence entre les méthodes . . . . .	13
1.6.1 PCA et son lien avec la KLD . . . . .	13
1.6.2 SVD et son lien avec la POD . . . . .	13
1.6.3 PCA et son lien avec la SVD . . . . .	14
1.7 Décomposition progressive généralisée (PGD) . . . . .	14
1.7.1 Produits tensoriels finis d'espace de Banach . . . . .	16
1.8 PGD progressive . . . . .	22
1.8.1 L'opérateur p-laplacien . . . . .	23
1.9 Plan de la thèse . . . . .	26
1.9.1 Résumé des chapitres . . . . .	26
<b>2 La sensibilité des éléments propres d'opérateurs compacts auto-adjoints et quelques applications</b>	<b>29</b>
2.1 Introduction . . . . .	30
2.2 Préliminaires et notations . . . . .	31
2.3 Régularité des valeurs propres et des vecteurs propres sur $\mathcal{K}_s(H)$ .	32
2.4 Application à la méthode POD . . . . .	37
2.4.1 Variation du paramètre . . . . .	38
2.4.2 La variation du domaine spatial : Gappy POD . . . . .	40
2.4.3 La variation de l'intervalle temporel . . . . .	41



<b>3</b>	<b>PGD et Flot optique</b>	<b>43</b>
	Présentation du chapitre . . . . .	44
3.1	Introduction . . . . .	44
3.2	Matériels et méthodes . . . . .	45
3.2.1	Approche PGD proposée pour les méthodes variationnelles	45
3.2.2	Implementation des approches PGD et Horn&Schunck en flot optique . . . . .	54
3.2.3	Simulations numériques . . . . .	58
3.2.4	Evaluation de la méthode PGD proposée . . . . .	59
3.3	Résultats . . . . .	59
3.4	Discussion . . . . .	61
3.5	Conclusion . . . . .	62
<b>4</b>	<b>PGD sur la variété de Nehari pour les problèmes concave-convexe</b>	<b>69</b>
	Présentation du chapitre . . . . .	70
4.1	Introduction . . . . .	70
4.2	Résultats préliminaires . . . . .	71
4.3	Résultats d'existence . . . . .	76
4.4	Decomposition Progressive Généralisée (PGD) . . . . .	77
4.5	Résultats numériques . . . . .	79
<b>5</b>	<b>Annexe</b>	<b>89</b>
5.1	Annexe . . . . .	90
5.1.1	Condition de Palais-Smale . . . . .	90
5.1.2	Principe Variationnel d'Ekeland . . . . .	90
5.1.3	Lemme de Brézis-Lieb . . . . .	91
5.1.4	Flot Optique . . . . .	91
	<b>Liste des travaux</b>	<b>95</b>
	<b>Bibliographie</b>	<b>97</b>

# Chapitre 1

## Introduction et présentation des travaux

### Sommaire

---

<b>1.1</b>	<b>Intoduction à la Décomposition par Projection Orthogonale (POD) définie sur un espace de Hilbert complexe . . . . .</b>	<b>2</b>
<b>1.2</b>	<b>Espaces de Hilbert complexes . . . . .</b>	<b>4</b>
<b>1.3</b>	<b>Opérateurs de Hilbert-Schmidt . . . . .</b>	<b>6</b>
<b>1.4</b>	<b>Intégrales de Bochner . . . . .</b>	<b>8</b>
<b>1.5</b>	<b>Décomposition par projection orthogonale POD . . .</b>	<b>9</b>
1.5.1	Méthode directe . . . . .	9
1.5.2	Méthode des snapshots . . . . .	12
<b>1.6</b>	<b>Equivalence entre les méthodes . . . . .</b>	<b>13</b>
1.6.1	PCA et son lien avec la KLD . . . . .	13
1.6.2	SVD et son lien avec la POD . . . . .	13
1.6.3	PCA et son lien avec la SVD . . . . .	14
<b>1.7</b>	<b>Décomposition progressive généralisée (PGD) . . . .</b>	<b>14</b>
1.7.1	Produits tensoriels finis d'espace de Banach . . . . .	16
<b>1.8</b>	<b>PGD progressive . . . . .</b>	<b>22</b>
1.8.1	L'opérateur p-laplacien . . . . .	23
<b>1.9</b>	<b>Plan de la thèse . . . . .</b>	<b>26</b>
1.9.1	Résumé des chapitres . . . . .	26

---

## 1.1 Introduction à la Décomposition par Projection Orthogonale (POD) définie sur un espace de Hilbert complexe

Pour tout problème physique, et à partir des informations nécessaires sur ce problème, on peut établir un modèle mathématique qui le représente. Cette modélisation nous donne une forme continue que l'on peut étudier mathématiquement pour chercher les différentes propriétés qui nous intéressent : existence des solutions, unicité, stabilité, comportement asymptotique..

Mises à part quelques équations académiques, on ne peut pas trouver la solution analytique de notre modèle. Cependant, on peut approximer le problème mathématique continu par un modèle numérique à l'aide d'une discrétisation adéquate qui rend l'erreur entre la solution théorique et la solution numérique petite.

Les problèmes à résoudre peuvent s'écrire sous la forme générique suivante :

$$\begin{cases} \frac{\partial u}{\partial t} - \mathcal{A}(u) &= f \\ u(x, t=0) &= u^0(x) \quad (\text{condition initiale}) \\ u(x \in \partial\Omega, t) &= u^b(t) \quad (\text{condition au bord}) \end{cases}$$

avec  $\mathcal{A}$ , l'opérateur spatial,  $\Omega = \Omega_x \times \Omega_t$  où  $\Omega_x$  est le domaine spatial et  $\Omega_t$  est le domaine temporel.

Etant donné que la précision du modèle numérique augmente avec la finesse de la discrétisation, les quantités de calcul deviennent très coûteuses pour une bonne approximation des solutions des problèmes complexes. Par exemple, en Interaction Fluide, si on remplace le modèle continu initial par un système d'équations obtenu par une discrétisation classique (différences finies, éléments finis, volumes finis,...), on remarque que notre problème contient un nombre énorme d'inconnus.

Ainsi, la plupart des problèmes importants ont des modèles complexes. Et on aura des difficultés à avoir les solutions en temps réel, à cause des fines discrétisations associées. Donc, afin d'avoir une solution en temps réel, la recherche de réduction de modèle a pris une place importante dans de nombreux domaines de la science et de l'ingénierie.

Pour réduire un modèle, il suffit de le projeter sur une base réduite de fonctions de dimension inférieure que la dimension du modèle numérique fin, associé au problème mathématique. Après cette projection, on obtient un système d'équations à très peu d'inconnues relativement au système d'équations obtenu par une discrétisation classique.

Maintenant, un bon modèle réduit, c'est celui qui conserve l'énergie de la solution de notre modèle continu, avec une approximation bien précise. Plus précisément, si notre problème dans le modèle continu est défini sur un espace vectoriel normé  $H$ , et si  $H_n$  est l'espace de dimension  $n$  dans laquelle la discrétisation avec les méthodes classiques de ce problème est défini, alors la réduction de modèle permet de remplacer l'un de deux problèmes (le problème continu ou le problème discret de dimension  $n$ ) par un problème de dimension  $N \ll n$ , à partir d'une projection sur un espace de dimension  $N$  appelé  $H_N$ . Une condition nécessaire pour choisir la bonne  $N$ , est la conservation de l'énergie du problème, c'est-à-dire :

$$||E(N) - E(n)|| \leq \epsilon \quad \text{où } \epsilon \text{ est la précision souhaitée}$$

Les méthodes les plus connues dans le domaine sont : la méthode POD introduite dans [73] et la méthode de décomposition généralisée PGD, introduite par Pierre Ladevèze [67] en (1985) où il a fait la séparation espace-temps, et qui a été appliquée à la résolution de problèmes multi-dimensionnels par Ammar, Chinesta et al. [32] en (2008). La méthode POD vise à trouver les premiers vecteurs propres dominants de la corrélation spatiale du champ de vitesse. Plus précisément, elle permet de représenter le champs de vitesse avec une base orthogonale et ordonnée optimale à partir d'une donnée  $u(t)$  connue sur un domaine temporel continu ou discret. L'efficacité de la POD vient de l'optimalité énergétique de la base retenue. Dans la littérature, il y a encore d'autres méthodes de réduction comme la méthode de troncature équilibrée [42], la décomposition de Hankel, et la méthode de Krylov définie par Alexei Krylov en (1931)...

À l'opposé de certaines bases utilisées en mathématiques comme les bases de Fourier, d'ondelettes, d'éléments finis ou de la PGD, qui permettent de définir *à priori* un espace de travail de dimension finie, pour résoudre un problème des Équations aux Dérivées Partielles (EDP) par exemple, la base POD n'est pas bien adaptée *à priori*, parce qu'elle est construite *à posteriori* à partir d'une référence  $u$ .

La POD a été introduite par Lumley en 1967 [73] dans le contexte de la turbulence. Elle a encore été introduite indépendamment dans d'autres disciplines, représentée par : Karhunen [52], Loève [71], Obukhov [85] et Kosambi [53]. C'est pour cela, elle porte plusieurs noms comme : la KLD (Décomposition de Karunen-loève), la PCA (Décomposition en Composantes Principales) [50], et la SVD (Décomposition en Valeurs Singulières) [13]. Dans [95], [70] et [97] les auteurs ont établi les connexions asymptotiques entre ces Karhunen [52], Loève [71], Obukhov [85] et Kosambi [53]. C'est pour cela, elle porte plusieurs noms comme : la KLD, la PCA [50], et la SVD [13]. Dans [95], [70] et [97], les auteurs ont établi les connexions asymptotiques entre ces méthodes et ils ont montré leur équivalence pour des problèmes de dimensions finies.

La POD a beaucoup d'applications pour la compression d'image [72], en contrôle optimal [88] pour les dynamiques des fluides : les écoulements dans des cavités [27], dans un canal ou autour d'un cylindre circulaire [34], les couches de cisaillement [87] et les transitions aux frontières des couches [89]. Cette méthode est utilisée dans de nombreux domaines (mécanique des fluides, aéronautique, vibration et systèmes chaotiques, thermique, etc..) pour construire des modèles avec des dimensions très petites, pour faire l'approximation de la solution du modèle principal, décrite par des équations aux dérivées partielles.

## 1.2 Espaces de Hilbert complexes

Dans tout ce qui suivra,  $H$  désignera un espace vectoriel complexe séparable.

**Définition 1.2.1.** (FORMES SESQUILINÉAIRES) *Soit  $\langle .; . \rangle$  une application de  $H \times H$  dans  $\mathbb{C}$ .*

*On dit que  $\langle .; . \rangle$  est une forme sesquilinéaire sur  $H$  ssi  $\langle .; . \rangle$  est linéaire par rapport à la première variable et sesquilinéaire par rapport à la deuxième variable, c'est-à-dire :*

*pour tout  $(x_1; y_1), (x_2; y_2)$  dans  $H \times H$  et pour tout  $(\lambda_1; \lambda_2), (\mu_1; \mu_2)$  dans  $\mathbb{C} \times \mathbb{C}$ , on a :*

$$\langle \lambda_1 x_1 + \lambda_2 x_2; y_1 \rangle = \lambda_1 \langle x_1; y_1 \rangle + \lambda_2 \langle x_2; y_1 \rangle$$

*et*

$$\langle x_1; \mu_1 y_1 + \mu_2 y_2 \rangle = \overline{\mu_1} \langle x_1; y_1 \rangle + \overline{\mu_2} \langle x_1; y_2 \rangle.$$

**Définition 1.2.2.** (FORME HERMITIENNE) *Une forme sesquilinéaire  $\langle .; . \rangle$  est dite hermitienne ssi pour tout  $(x; y) \in H \times H$ ;  $\langle x; y \rangle = \overline{\langle y; x \rangle}$ .*

**Définition 1.2.3.** (PRODUIT SCALAIRE) *On appelle  $\langle .; . \rangle$  produit scalaire sur  $H$  si :*

1.  $\langle .; . \rangle$  est une forme sesquilinéaire Hermitienne.
2.  $\forall x \in H, \langle x; x \rangle \geq 0$ .
3.  $\forall x \in H, \langle x; x \rangle = 0 \Rightarrow x = 0_H$ .

**Définition 1.2.4.** *On appelle espace préhilbertien complexe tout  $\mathbb{C}$ -espace vectoriel muni d'un produit scalaire.*

**Proposition 1.2.5.** (INÉGALITÉ DE CAUCHY-SCHWARTZ) *Soient  $x$  et  $y$ , 2 éléments non nuls dans  $H$  qui est un espace pré-hilbertien, et on a :*

1.  $|\langle x, y \rangle|^2 \leq \langle x, x \rangle \langle y, y \rangle$ .

2.  $|\langle x, y \rangle|^2 = \langle x, x \rangle \langle y, y \rangle$  si et seulement si  $x$  et  $y$  sont linéairement dépendants.

**Preuve.**

1. Soient  $x$  et  $y$  non nuls dans  $H$ , alors pour tout  $\lambda \in \mathbb{R}$  on a :

$$\begin{aligned} 0 &\leq \langle x + \lambda y, x + \lambda y \rangle = \langle x, x \rangle + \lambda \langle y, x \rangle + \lambda \langle x, y \rangle + \lambda^2 \langle y, y \rangle \\ &= \langle x, x \rangle + 2\lambda \Re(\langle x, y \rangle) + \lambda^2 \langle y, y \rangle. \\ &\leq \langle x, x \rangle + 2\lambda |\langle x, y \rangle| + \lambda^2 \langle y, y \rangle. \end{aligned}$$

Ainsi, pour tout  $\lambda \in \mathbb{R}$ , on a :

$$\langle x, x \rangle + 2\lambda |\langle y, x \rangle| + \lambda^2 \langle y, y \rangle \geq 0.$$

Par conséquent, le discriminant du précédent trinôme est négatif, c'est-à-dire :

$$|\langle x, y \rangle|^2 - \langle x, x \rangle \langle y, y \rangle \leq 0.$$

D'où l'inégalité de Cauchy-Schwartz.

2. Si on a égalité, alors le discriminant précédent est nul, donc il existe un  $\lambda_0 \in \mathbb{R}$  tel que  $\langle x + \lambda_0 y, x + \lambda_0 y \rangle = 0$ , ou encore  $x + \lambda_0 y = 0$ , d'où le résultat.

□

**Proposition 1.2.6.** (NORME HERMITIENNE) Soit  $H$  un espace préhilbertien alors  $x \mapsto \|\cdot\| := \sqrt{\langle \cdot, \cdot \rangle}$  est une norme sur  $H$ , appelée norme hermitienne.

**Définition 1.2.7.** (ESPACE DE HILBERT) Un espace préhilbertien  $(H, \langle \cdot, \cdot \rangle)$ , complet pour la norme  $\|\cdot\| = \sqrt{\langle \cdot, \cdot \rangle}$  est appelé un **espace de Hilbert**.

**Lemme 1.2.8.** (INÉGALITÉ DE BESSEL) Considérons une suite orthonormée  $(e_n)_{n \geq 0}$  dans un espace de Hilbert  $H$ , pour tout  $x \in H$ , la série numérique est convergente et :

$$\sum_{n=0}^{\infty} |\langle x, e_n \rangle|^2 \leq \|x\|^2.$$

**Définition 1.2.9.** (BASE HILBERTIENNE) Une base hilbertienne d'un Hilbert  $H$  est une suite d'éléments  $(e_n)_{n \geq 0}$  de  $H$  telle que :

1. Pour tous  $n, m \in \mathbb{N}$ ,  $\langle e_n, e_m \rangle = \delta_{m,n}$ .
2. La famille formée par les  $(e_n)$  est totale, c'est-à-dire : l'espace vectoriel engendré par la famille  $(e_n)_{n \geq 0}$  est dense dans  $H$ .

Dans ce cas, pour tout  $x \in H$ , on a :

$$x = \sum_{n=0}^{\infty} \langle x, e_n \rangle e_n \quad \text{et} \quad \|x\|^2 = \sum_{n=0}^{\infty} |\langle x, e_n \rangle|^2 \quad (\text{égalité de Parseval}).$$

**Proposition 1.2.10.** *Tout espace de Hilbert séparable admet une base hilbertienne.*

### 1.3 Opérateurs de Hilbert-Schmidt

**Proposition 1.3.1.** *Soient  $H_1$  et  $H_2$  des espaces de Hilbert et  $K \in \mathcal{L}(H_1; H_2)$ . Alors il existe un unique  $K^* \in \mathcal{L}(H_2; H_1)$  tel que, pour tout  $x \in H_1$  et tout  $y \in H_2$ , on ait :*

$$\langle Kx, y \rangle = \langle x, K^*y \rangle.$$

**Remarque 1.3.2.** *Avec les mêmes notations de la proposition (1.3.1), on appelle  $K^*$  l'adjoint de l'opérateur  $K$ .*

**Proposition 1.3.3.** *Soient  $H_1, H_2$  des espaces des Hilbert,  $K \in \mathcal{L}(H_1; H_2)$ . Alors :*

1.  $K^*$  est borné et  $\|K\| = \|K^*\|$ ;
2.  $(K^*)^* = K$ ;
3. Si  $K$  est inversible,  $(K^{-1})^* = (K^*)^{-1}$ .

Dans toute la suite, l'espace dual  $\mathcal{L}(H, \mathbb{R})$  sera muni de sa norme naturelle, induite par celle de  $H$ . Cette norme duale sera notée également par  $\| \cdot \|$ .

**Définition 1.3.4.** *Soit  $H$  un espace de Hilbert, un opérateur  $K \in \mathcal{L}(H)$  est dit **auto-adjoint** si  $K^* = K$ .*

**Définition 1.3.5.** *L'opérateur  $K \in \mathcal{L}(H)$  est dit de Hilbert-Schmidt s'il existe une base Hilbertienne  $(e_n)_{n \in \mathbb{N}}$  de  $H$  telle que :*

$$\sum_{n \in \mathbb{N}} \|K e_n\|_H^2 < +\infty.$$

**Théorème 1.3.6.** *Tout Opérateur Hilbert-Schmidt est compact.*

**Preuve.** Soit  $(e_n)$  une base hilbertienne de  $H$  et soit  $K$  un opérateur Hilbert-Schmidt sur  $H$ . Pour tout  $n$ , on désigne par  $\Pi_n$  l'opérateur de projection orthogonale sur le sous-espace engendré par  $\{e_1, e_2, \dots, e_n\}$ . L'opérateur  $K_n = K \circ \Pi_n$  est de rang fini égal à  $n$  et on a :

$$K_n e_j = \begin{cases} K e_j, & \text{si } j \leq n \\ 0, & \text{si } j \geq n + 1 \end{cases}$$

donc :

$$(K - K_n) e_j = \begin{cases} 0, & \text{si } j \leq n \\ K e_j, & \text{si } j \geq n + 1. \end{cases}$$

Il en résulte que pour tout  $x \in H$ , on a :

$$(K - K_n) x = \sum_{j=n+1}^{+\infty} \langle x; e_j \rangle K e_j.$$

Ainsi :

$$\begin{aligned} \|(K - K_n) x\| &\leq \sum_{j=n+1}^{+\infty} |\langle x; e_j \rangle| \times \|K e_j\| \\ &\leq \left( \sum_{j=n+1}^{+\infty} |\langle x; e_j \rangle|^2 \right)^{\frac{1}{2}} \times \left( \sum_{j=n+1}^{+\infty} \|K e_j\|^2 \right)^{\frac{1}{2}} \\ &\leq \|x\|^2 \times \left( \sum_{j=n+1}^{+\infty} \|K e_j\|^2 \right)^{\frac{1}{2}}. \end{aligned}$$

D'où :

$$\|(K - K_n)\| \leq \left( \sum_{j=n+1}^{+\infty} \|K e_j\|^2 \right)^{\frac{1}{2}},$$

or le second membre de cette inégalité tend vers zéro lorsque  $n$  tend vers l'infini, puisque c'est le reste d'une série convergente. L'opérateur  $K$  est donc une limite d'une suite d'opérateurs de rang fini.  $\square$

**Théorème 1.3.7.** *Si  $H$  est un espace de Hilbert séparable et  $K$  un opérateur compact auto-adjoint sur  $H$  alors,  $H$  admet une base hilbertienne formée de vecteurs propres de  $K$ . Donc il existe une base orthonormale complète  $\{\varphi_i\}_{i=1}^{+\infty}$  telle que :*

$$K \varphi_i = \lambda_i \varphi_i$$

où  $\{\lambda_i\}_{i=1}^{+\infty}$ , une suite décroissante de valeurs propres, à valeurs positives

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n \geq \dots 0$$

telle que  $\lambda_n \rightarrow 0$  lorsque  $n \rightarrow \infty$  et, pour tout  $u \in H$ , on a :  $u = \sum_{i=1}^{+\infty} \langle u, \varphi_i \rangle_H \varphi_i$ .



## 1.4 Intégrales de Bochner

**Définition 1.4.1.** (ESPACE MESURABLE) *Un couple  $(X, \mathcal{A})$  constitué d'un ensemble non vide  $X$  et d'une tribu  $\mathcal{A}$ , formée de sous ensembles de  $X$ , s'appelle un espace mesurable. Les éléments de  $\mathcal{A}$  sont les ensembles mesurables de  $X$ .*

On présente succinctement la généralisation de l'intégrale de Lebesgue pour les fonctions définies sur  $[0, T]$  et à valeurs dans  $H$ .

**Définition 1.4.2.** (INTÉGRALE DE BOCHNER) *Une fonction  $f : [0, T] \longrightarrow H$  est dite étagée s'il existe une famille finie de parties mesurables  $(I_k)_k \subset [0, T]$  (pour la mesure de Lebesgue) et de vecteurs  $(a_k)_k \subset H$ , avec  $1 \leq k \leq n$ , telle que :*

$$f = \sum_{k=1}^n a_k \chi_{I_k}, \quad \text{où } \chi_{I_k} \text{ désigne la fonction indicatrice de l'ensemble } I_k.$$

Dans ce cas, on définit l'intégrale de  $f$  par :

$$\int_0^T f(t) dt := \sum_{k=1}^n a_k \mu(I_k), \quad \text{où } \mu \text{ désigne la mesure de Lebesgue standard sur } \mathbb{R}.$$

Une fonction mesurable  $f : [0, T] \longrightarrow H$  est dite intégrable sur  $[0, T]$ , au sens de Bochner, s'il existe une suite de fonctions étagées  $f_n : [0, T] \longrightarrow H$ ,  $n \in \mathbb{N}$ , telle que :

$$\lim_{n \rightarrow +\infty} \int_0^T \|f_n(t) - f(t)\|_H dt = 0.$$

Dans ce cas, on définit l'intégrale de Bochner de  $f$  sur  $[0, T]$  par :

$$\int_0^T f(t) dt := \lim_{n \rightarrow +\infty} \int_0^T f_n(t) dt.$$

**Définition 1.4.3.** (FONCTION MESURABLE) *Soient  $(X, \mathcal{A})$  et  $(Y, \mathcal{B})$  deux espaces mesurables. Une fonction  $f : X \longrightarrow Y$  est dite mesurable si*

$$\forall B \in \mathcal{B}, \quad f^{-1}(B) \in \mathcal{A}.$$

**Proposition 1.4.4.** *Une fonction mesurable  $f : [0, T] \longrightarrow H$  est intégrable au sens de Bochner si et seulement si :*

$$\int_0^T \|f(t)\|_H dt < +\infty.$$

De plus, si  $A : H \longrightarrow \mathbb{R}$  est un opérateur linéaire continu (ou même fermé), alors :

$$A \left( \int_0^T f(t) dt \right) = \int_0^T A(f(t)) dt$$

## 1.5 Décomposition par projection orthogonale POD

La POD permet de définir une base orthogonale et ordonnée optimale pour représenter  $u(x, t)$  à partir d'une  $u(x, t)$  connue sur un domaine temporel  $\mathcal{T}$ . Cette base nous permet d'identifier les degrés de libertés les plus représentatifs pour construire un modèle d'ordre réduit.

Donc, le but de la méthode POD est d'approcher une fonction  $u(x, t)$ , par des fonctions à variables séparées du type :

$$u(x, t) \simeq \sum_{i=1}^N a_i(t) \varphi_i(x)$$

où  $N$  est fini,  $x$  et  $t$  sont les variables d'espace et de temps, respectivement.

### 1.5.1 Méthode directe

Généralement, on peut résoudre ce problème *à priori* en considérant les  $\varphi_i$  comme des fonctions connues, par exemple, les polynômes de Chebychev ou de Legendre. Par contre, ici, on va chercher les  $\varphi_i$  les plus énergétiques de la fonction  $u$  à approximer, ce qui fait de la POD une approximation *à posteriori*.

Soit  $u \in L^2([0, T]; H)$  une fonction fixée et soit  $K$  l'opérateur défini par :

$$K : H \longrightarrow H$$

$$\varphi \longmapsto K\varphi := \frac{1}{T} \int_0^T \langle \varphi, u(t) \rangle_H \times u(t) dt.$$

Notons que l'opérateur linéaire  $K$  est bien défini, c'est-à-dire que la fonction  $t \mapsto \langle \varphi, u(t) \rangle_H u(t)$  est bien dans  $L^1([0, T]; H)$ .

**Proposition 1.5.1.** *L'opérateur  $K$  est un opérateur de Hilbert-Schmidt (donc compact) et auto-adjoint.*

#### Preuve

- $K$  est compact.

Soit  $(e_n)_{n \in \mathbb{N}}$  une base Hilbertienne de  $H$  quelconque. On a alors :

$$\begin{aligned}
 T^2 \sum_n \|K e_n\|_H^2 &= \sum_n \left\| \int_0^T \langle e_n, u(t) \rangle_H u(t) dt \right\|^2 \\
 &\leq \sum_n \left( \int_0^T |\langle e_n, u(t) \rangle_H| \times \|u(t)\| dt \right)^2 \\
 &\leq \sum_n \left( \int_0^T |\langle e_n, u(t) \rangle_H|^2 dt \right) \|u(t)\|_{L^2([0,T];H)}^2 \\
 &= \|u(t)\|_{L^2([0,T];H)}^2 \int_0^T \sum_n |\langle e_n, u(t) \rangle_H|^2 dt \\
 &= \|u(t)\|_{L^2([0,T];H)}^4.
 \end{aligned}$$

–  $K$  est auto-adjoint.

On utilisera la continuité du produit scalaire de  $H$  dans  $\mathbb{C}$  ainsi que les propriétés élémentaires du produit scalaire. En effet,

$$\begin{aligned}
 \langle K \varphi; \psi \rangle &= \frac{1}{T} \left\langle \int_0^T \langle \varphi, u(t) \rangle_H u(t) dt; \psi \right\rangle_H \\
 &= \frac{1}{T} \int_0^T \langle \varphi, u(t) \rangle_H \langle u(t), \psi \rangle_H dt \\
 &= \langle \varphi; K \psi \rangle_H.
 \end{aligned}$$

D'où le résultat. □

Soit  $u \in L^2([0, T]; H)$  une fonction donnée où  $T > 0$  est fixé. Notre but est de chercher la base Hilbertienne de  $H$  qui soit la plus proche de la moyenne de  $u$  sur  $[0, T]$ . Pour cela, considérons le problème d'optimisation :

$$\sup_{\varphi \in H \setminus \{0_H\}} \frac{\frac{1}{T} \int_0^T |\langle \varphi, u(t) \rangle_H|^2 dt}{\|\varphi\|_H^2} \quad (1.1)$$

qui consiste à la recherche de la direction (ou les directions) la plus proche de  $u$  dans l'espace  $H$ , au sens de la moyenne de  $u$  sur  $[0, T]$ .

Posons :

$$J(\varphi) := \frac{1}{T} \int_0^T |\langle \varphi, u(t) \rangle_H|^2 dt.$$

**Théorème 1.5.2.** *Pour tout  $\varphi \in H$ , on a :*

$$J(\varphi) = \langle \varphi, K \varphi \rangle_H.$$

La preuve est immédiate. Elle utilise la continuité du produit scalaire.

**Théorème 1.5.3.** (1.1) a une solution unique.

**Preuve.** Fixons  $\varphi \in H \setminus \{0_H\}$ . Soit  $0 < \epsilon \ll 1$ , avec  $\epsilon < \|\varphi\|$  et soit  $\psi \in \mathcal{B}(0, \epsilon)$ .

On définit :

$$\begin{aligned} F : H / \{0_H\} &\longrightarrow \mathbb{R} \\ \varphi &\longmapsto F(\varphi) = \frac{J(\varphi)}{\langle \varphi, \varphi \rangle_H}. \end{aligned}$$

$$\begin{aligned} F(\varphi + \psi) &= \frac{J(\varphi + \psi)}{\langle \varphi + \psi, \varphi + \psi \rangle_H} = \frac{\langle \varphi + \psi, K(\varphi + \psi) \rangle_H}{\langle \varphi + \psi, \varphi + \psi \rangle_H} \\ &= \frac{\langle \varphi, K\varphi \rangle_H + \langle \psi, K\varphi \rangle_H + \langle \varphi, K\psi \rangle_H + \langle \psi, K\psi \rangle_H}{\langle \varphi, \varphi \rangle_H + \langle \varphi, \psi \rangle_H + \langle \psi, \varphi \rangle_H + \langle \psi, \psi \rangle_H} \\ &= \frac{\langle \varphi, K\varphi \rangle_H + \langle \psi, K\varphi \rangle_H + \langle \varphi, K\psi \rangle_H + \theta(\|\psi\|_H^2)}{\langle \varphi, \varphi \rangle_H + \langle \varphi, \psi \rangle_H + \langle \psi, \varphi \rangle_H + \theta(\|\psi\|_H^2)} \\ &= \frac{\langle \varphi, K\varphi \rangle_H + 2\Re\langle \psi, K\varphi \rangle_H + \theta(\|\psi\|_H^2)}{\langle \varphi, \varphi \rangle_H + 2\Re\langle \psi, \varphi \rangle_H + \theta(\|\psi\|_H^2)} \\ &= \frac{F(\varphi) + \frac{2\Re\langle \psi, K\varphi \rangle_H}{\langle \varphi, \varphi \rangle_H} + \theta(\|\psi\|_H^2)}{1 + \frac{2\Re\langle \psi, \varphi \rangle_H}{\langle \varphi, \varphi \rangle_H} + \theta(\|\psi\|_H^2)} \\ &= \left( F(\varphi) + \frac{2\Re\langle \psi, K\varphi \rangle_H}{\langle \varphi, \varphi \rangle_H} + \theta(\|\psi\|_H^2) \right) \left( 1 - \frac{2\Re\langle \psi, \varphi \rangle_H}{\langle \varphi, \varphi \rangle_H} - \theta(\|\psi\|_H^2) \right) \\ &= F(\varphi) + \frac{2\Re\langle \psi, K\varphi \rangle_H}{\langle \varphi, \varphi \rangle_H} - 2F(\varphi) \frac{\Re\langle \psi, \varphi \rangle_H}{\langle \varphi, \varphi \rangle_H} + \theta(\|\psi\|_H^2) \end{aligned}$$

Si  $\varphi$  réalise le maximum dans (1.1), alors :

$$\frac{2\Re\langle \psi, K\varphi \rangle_H}{\langle \varphi, \varphi \rangle_H} - 2F(\varphi) \frac{\Re\langle \psi, \varphi \rangle_H}{\langle \varphi, \varphi \rangle_H} = 0$$

et donc on obtient :

$$\Re(\langle \psi, K\varphi \rangle_H) = F(\varphi) \Re(\langle \psi, \varphi \rangle_H).$$

En refaisant les mêmes calculs avec  $i\psi$  à la place de  $\psi$ , on obtient :

$$\Re(\langle i\psi, K\varphi \rangle_H) = F(\varphi) \Re(\langle i\psi, \varphi \rangle_H).$$

c'est-à-dire :

$$\text{Im}(\langle \psi, K\varphi \rangle_H) = F(\varphi) \text{Im}(\langle \psi, \varphi \rangle_H).$$

D'où :

$$K\varphi = \lambda\varphi, \quad (1.2)$$

$\forall \psi \in \mathcal{B}(0, \epsilon)$  où  $\lambda = F(\varphi)$ .

Réciproquement, tout vecteur propre  $\varphi$  correspondant à la plus grande valeur propre  $\lambda_1$  de  $K$ , satisfait (1.1) et donc notre problème s'agit d'un problème aux valeurs propres.  $\square$

En utilisant le théorème (1.3.7), nous déduisons l'existence d'une base ortho-normale complète  $\{\varphi_k\}_{k=1}^{+\infty}$  telle que pour tout  $u \in H$  on a :  $u = \sum_{i=1}^{+\infty} \langle u, \varphi_i \rangle_H \varphi_i$ .

**Définition 1.5.4.** Soit  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n \geq \dots 0$ , la suite décroissante des valeurs propres de l'opérateur  $K$ . Pour tout  $N \in \mathbb{N}$ , on appelle **base POD d'ordre  $N$** , toute famille  $\{\varphi_k\}_{k=1}^N$  de vecteurs propres associés, et on a :

$$\forall i, j \in \{1, \dots, N\}; \quad \langle \varphi_i, \varphi_j \rangle_H = \delta_{i,j}$$

où  $\delta$  est l'opérateur de kronecker.

**Remarque 1.5.5.** Dans la méthode POD directe, les fonctions propres  $\{\varphi_k\}_{k=1}^{\infty}$  sont purement spatiales.

## 1.5.2 Méthode des snapshots

En 1987, Sirovich a introduit une méthode symétrique à la POD classique [93], que l'on appelle méthode de snapshot. Cette méthode consiste à chercher une base à l'aide d'un moyen temporel et non pas spatial. Donc les éléments de base de cette méthode ne dépendent que du temps, à l'opposé du cas de la méthode classique.

Si l'écoulement peut se représenter par  $M$  clichés et  $M \ll d.n$  où  $d$  est la dimension de l'espace et  $n$  est le nombre de noeuds, alors la taille du problème de snapshot est  $M^2$ , beaucoup plus petit, par rapport à la taille de la méthode classique qui est égale à  $(n.d)^2$ .

Donc cette méthode est moins coûteuse pour les mémoires calculatoires lorsque le nombre d'échantillons temporels est beaucoup plus petit que le nombre de points en espace. Par exemple, on a le cas des écoulements inhomogènes [43].

Le principe de la méthode consiste à écrire chaque vecteur propre de l'opérateur  $K$  comme une combinaison linéaire des champs  $u(x, t_i)$  :

$$\varphi_i(x) = \sum_{j=1}^{N_t} a_i(t_j) u(x, t_j), \quad \text{où } N_t \text{ est le nombre des échantillons temporels.}$$

Maintenant, on met les  $\varphi_i$  par leurs valeurs dans l'équation (1.2) pour trouver les  $a_i(t_j)$ ,  $\forall j = 1, \dots, N_t$ .

On aura :

$$R_{j\,k}.a_i(t_k) = \lambda_i a_i(t_k), \quad \forall k = 1, \dots, N_t$$

où  $R_{j\,k} = \frac{1}{N_t} \int_{\Omega} u(t_j, x) u(t_k, x) dx$ , et  $\lambda_i$  est la valeur propre associée à la mode  $\varphi_i$  de la POD classique.

## 1.6 Equivalence entre les méthodes

La POD a été introduite par plusieurs scientifiques et pour des utilisations différentes ; par exemple, par Pearson pour l'analyse des données, par Lumley dans le domaine de la mécanique des fluides et par Karhunen et Loève en imagerie. C'est pour cette raison qu'elle est connue sous plusieurs noms : POD, KLD, PCA et SVD. Des relations directes relient ces méthodes pour les considérer comme une seule méthode avec des représentations différentes.

### 1.6.1 PCA et son lien avec la KLD

Dans le domaine de traitement des images, Karhunen et Loève ont développé indépendamment la décomposition KLD. On peut considérer cette décomposition comme une PCA dans les espaces des dimensions infinies. Dans [95], les auteurs ont utilisé le théorème de Mercer et le théorème de Diamantaras et Kung pour démontrer l'équivalence entre ces deux méthodes.

### 1.6.2 SVD et son lien avec la POD

La SVD permet d'approximer un système de dimension élevée par un autre système de dimension beaucoup plus faible, à l'aide du théorème fondamental de décomposition en algèbre linéaire.

Soit  $A$  une matrice complexe de taille  $M \times N$ , alors la décomposition SVD de  $A$  est la factorisation :

$$A = U \Sigma V^T$$

où  $U$  et  $V$  sont deux matrices unitaires de dimensions  $M \times M$ ,  $N \times N$  respectivement, c'est-à-dire :

$$U U^T = I_M \quad \text{et} \quad V V^T = I_N,$$

et  $\Sigma = \text{diag}(\sigma_1, \dots, \sigma_r)$  est une matrice diagonale de dimension  $M \times N$ , où les  $\sigma_i$  sont les valeurs singulières de  $A$  telles que,  $\sigma_1 \geq \dots \geq \sigma_r > \sigma_{r+1} = \sigma_{r+2} = \dots = \sigma_p = 0$  où  $p = \min(M, N)$ . Le nombre de valeurs singulières non-nulles de la matrice  $A$  est égale à  $p$ , le rang de cette matrice. On appelle les  $p$  premières

colonnes de  $V$  et de  $U$  les vecteurs singuliers droit et gauche respectivement de  $A$ . Puisque  $U$  et  $V$  sont deux vecteurs orthogonaux, alors :

$$A^T.A = V\Sigma^2V^T \text{ et } A.A^T = U\Sigma^2U^T.$$

Or les deux matrices  $A^T.A$  et  $A.A^T$  sont symétriques et définies positives donc diagonalisables dans des bases orthonormées de vecteurs propres, et leurs valeurs propres sont positives. Alors il existe  $W$  tel que :

$$A^T.A = W\Lambda W^T \text{ et } A.A^T = W\Lambda W^T.$$

D'après ce qui précède, on trouve  $\Sigma^2 = \Lambda$  et  $W = V$  dans le premier cas, et  $W = U$  pour le cas  $A.A^T$ .

Enfin, on peut conclure que,  $(U, \Lambda)$  représentent la POD de  $A.A^T$  et  $(V, \Lambda)$  représentent la POD de  $A^T.A$ , et les valeurs singulières  $\sigma_i$  de  $A$  sont les racines carrées des valeurs propres  $\lambda_i$  de  $A.A^T$ . Pour obtenir l'expression de la POD de  $A = U\Sigma V^T$ , on considère les colonnes de la matrice  $U\Sigma$  en tant que les fonctions de temps, et les colonnes de la matrice  $V^T$  comme les fonctions spatiales.

### 1.6.3 PCA et son lien avec la SVD

La PCA a été proposée par Karl Pearson en 1901. Elle consiste à réduire la dimension d'un ensemble de données corréliées en retenant les  $n$  composantes principales non corréliées de cet ensemble, où  $n$  est très petit par rapport à la dimension de l'espace initiale. Ces  $n$  composantes peuvent conserver l'énergie du système initial.

En SVD, il faut étudier la matrice transposée  $X^T$  de la matrice  $X$  qui concerne notre système. Les valeurs singulières obtenues sont les racines carrées arithmétiques des valeurs propres de la matrice  $X.X^T$ . Les vecteurs singuliers, choisis comme base, sont aussi les vecteurs propres de la matrice  $X.X^T$ .

Pour montrer l'équivalence de problème de valeurs propres entre SVD et PCA, il faut étudier la matrice  $\frac{1}{\sqrt{n}}X^T$  à la place de  $X^T$ . Pour exécuter la SVD de la nouvelle matrice, il faut étudier le problème des valeurs propres et des vecteurs propres de la matrice  $\frac{1}{n}X.X^T$  qui est bien équivalent au problème des valeurs et des vecteurs propres de  $X.X^T$ . En [97], les auteurs ont démontré que la matrice  $\frac{1}{n}X.X^T$  nous donne le problème des valeurs propres des deux méthodes PCA et KLD. Donc les vecteurs de la base obtenus, en utilisant le SVD au  $X^T$ , sont les mêmes que ceux qui sont obtenus en utilisant la PCA. De plus, en [35], Diamataras et Kung ont démontré une connection asymptotique entre les deux méthodes.

## 1.7 Décomposition progressive généralisée (PGD)

La méthode PGD est l'une des méthodes de réduction de modèles très connues. Elle a été développée par [10], [11], [66], pour trouver une solution approximative

d'une EDP d'une façon progressive et sous la forme de somme de produits de fonctions à variables séparées.

Contrairement à la méthode POD, la PGD est de type *à priori*. Elle consiste à construire, sans avoir des informations initiales, la solution approchée  $u$  dans un espace tensoriel  $V = V_1 \otimes \cdots \otimes V_d$ , où les  $V_i$  sont des espaces vectoriels de dimensions  $n_i$  pour tout  $i \in \{1, \dots, d\}$ . Lorsque  $m$  est suffisamment grande, on a :

$$u \approx u_m = \sum_{i=1}^m w_i^1 \otimes \cdots \otimes w_i^d$$

où  $w_i^j \in V_j$ . Dans [62], [10] et [79], une construction de la suite  $u_m$  est introduite d'une manière progressive, c'est-à-dire à chaque itération, on doit chercher un tenseur  $z$  de longueur 1 tel que :

$$u_m = u_{m-1} + z_m,$$

où  $z_m = \otimes_{k=1}^d w_m^k$ .

En d'autres termes, on doit écrire la solution sous la forme d'une somme des produits de  $m$  fonctions unidimensionnelles :

$$u(x_1, \dots, x_m) = \sum_{i=1}^m f_1^i(x_1) f_2^i(x_2) \cdots f_m^i(x_m).$$

Pour écrire les choses d'une façon claire, introduisons l'équation différentielle générale :

$$\mathcal{A}(u) = f,$$

où  $\mathcal{A}(u)$  est un opérateur différentiel qui dépend de  $u$ . La solution PGD  $u$  s'écrit sous une forme séparée qui dépend de  $m$  variables spatiales ou temporelles et de  $d$  fonctions à variables séparées telle que :

$$u(x_1, x_2, \dots, x_m) = \sum_{i=1}^d \prod_{j=1}^m F_j^i(x_j) = \sum_{i=1}^d F_1^i(x_1) F_2^i(x_2) \dots F_m^i(x_m).$$

Grâce à cette séparation des variables, une réduction impressionnante de calcul est obtenue. Ammar a montré dans [10] que la PGD fournit un problème de dimension  $10^5$  à la place d'un problème de dimension  $10^{40}$  avec la méthode des éléments finis. Plus généralement, on peut remplacer un problème résolu par les méthodes classiques de dimensions  $n^d$  noeuds, par un problème qui a besoin juste de  $n \times d$  noeuds après cette séparation.

Les domaines d'applications de la PGD sont trop nombreux. Nous allons présenter ici quelques travaux qui utilisent cette méthode dans des domaines différents. La PGD a été introduite par ladevèze [62] pour réduire les coûts de calcul dans la résolution du problème global linéaire de la méthode LATIN



(LArge Time INcrement method) introduite en 1985 [63], [64]. Donc, on peut voir la méthode PGD comme une méthode LATIN généralisée. Avant l'introduction de la PGD sous sa forme actuelle, plusieurs améliorations ont été effectuées sur la méthodes LATIN, dans le domaine des grands déplacements [23], [59]. Dans [20], [21], [28], [57], [58], une technique de décomposition de domaine a été proposée, et une caractéristique multi-échelles dans l'espace a été introduite dans [60], [61] et [65].

Une nouvelle stratégie multi-paramétriques a été présentée dans [4], [23], [24], [25] et [26]. Elle est appliquée pour les problèmes multi-physiques dans [37], [76], [77]. A. Nouy et P. Ladevèze ont appliqué la PGD couplée avec la méthode LATIN et ils ont trouvé des résultats pertinents au niveau du stockage et du temps de calcul [80], [81]. La PGD est aussi appliquée pour calculer les EDP stochastiques dans [82], [83] et [84]. Ainsi la PGD a été appliquée dans le domaine des problèmes multi-dimensionnels, par Ammar et al. dans [10], [11] et [12], afin de rendre le problème beaucoup plus léger, dans le domaine de séparation de l'espace physique dans de [36] et [45], et pour la simulation en temps réel dans [44], [69] et [78]. Pour plus de détails autour de la PGD et ses applications, les travaux de Chinesta et al. [29], [30], [31], peuvent donner des informations complémentaires.

### 1.7.1 Produits tensoriels finis d'espace de Banach

D'abord, nous rappelons quelques propriétés classiques des espaces tensoriels de Banach.

#### Définition 1.7.1. (PRODUIT TENSORIEL ALGÈBRIQUE)

Soit  $\{E_i, 1 \leq i \leq d\}$  l'ensemble des espaces  $E_i$ . Pour tout  $i = 1, \dots, d$ ,  $E_i$  est un espace de Banach muni d'un produit scalaire  $\|\cdot\|$ , où  $i$  signifie la  $i$ ème direction ou la  $i$ ème axe d'une fonction  $f(x_1, \dots, x_d)$ . On définit l'espace tensoriel algébrique  $E$  par :

$$E = a \bigotimes_{i=1}^d E_i,$$

où le suffixe  $a$  signifie la nature algébrique du produit tensoriel et  $d$  signifie la dimension spéciale. Par définition, dans un produit tensoriel algébrique, chaque élément de  $E$  est formé par les combinaisons linéaires finies des tenseurs primaires  $v = \bigotimes_{i=1}^d v^i$  où  $v^i \in E_i$ .

La représentation de Tucker est l'écriture typique qui a le forme :

$$u = \sum_{i=1}^d a_i \bigotimes b_{i_j}^j, \quad (1.3)$$

où  $I = I_1 \times \dots \times I_n$  est l'ensemble d'indices multiples avec  $I_j = \{1, \dots, r_j\}$ ,  $r_j \leq \dim(E_j)$ , les  $b_{i_j}^j \in E_j$  ( $i_j \in I_j$ ) sont orthonormales et linéairement indépendantes

et les  $a_i \in \mathcal{R}$ . Les éléments de  $I$  sont les  $d$  uplets et les  $i_j$  sont les composantes de  $i = (i_1, \dots, i_d)$ . Un tenseur  $V$  est défini par ses composantes  $V_i$ , tel que :

$$V := (V[i])_{i \in I}, \text{ où } V_i = V[i] = V[i_1, \dots, i_d] \in \mathbb{R}.$$

La taille des données est déterminée par les nombres  $r_j$  rassemblés dans  $r = (r_1, \dots, r_d)$ . Si on fixe  $r$ , l'ensemble de tous les tenseurs représentés par (1.7.7) est :

$$\mathcal{T}_r(E) := \left\{ v \in E : \begin{array}{l} \text{il existe des sous-espaces } U_j \subset E_j \\ \text{tels que } \dim(U_j) = r_j \text{ et } v \in U := {}_a \bigotimes_{i=1}^d U_i \end{array} \right\}, \quad (1.4)$$

et la lettre  $\mathcal{T}_r$  signifie la forme Tucker, ou le mot tenseur.

**Remarque 1.7.2.** Si  $r_j = 0$ ,  $\forall j = 1, \dots, d$ , alors dans ce cas là,  $\mathcal{T}_r = \mathcal{R}_0 = \{0\}$  qui est le sous-espace trivial.

Dans ce qui suit, on va travailler avec les tenseurs élémentaires de rang 1 tels que :

$$\mathcal{R}_1(V) := \mathcal{T}_{(1, \dots, 1)}(V) = \left\{ \bigotimes_{i=1}^d w^i, w^i \in V_i \right\},$$

et on définit encore  $\mathcal{R}_m$  tel que :

$$\mathcal{R}_m(V) := \left\{ \sum_{i=1}^m z_i, z_i \in \mathcal{R}_1(V) \right\}.$$

On dit que  $V_{\|\cdot\|}$  est un espace de Banach de tenseur s'il existe un espace de tenseur algébrique  $V$  muni de la norme  $\|\cdot\|$ , tel que  $V_{\|\cdot\|}$  soit la complétude de  $V$  par rapport à  $\|\cdot\|$ , donc :

$$V_{\|\cdot\|} := \overline{\|\cdot\| \bigotimes_{j=1}^d V_j} = \overline{{}_a \bigotimes_{j=1}^d V_j}^{\|\cdot\|}.$$

Si  $V_{\|\cdot\|}$  est un espace de Hilbert, on dit que  $V_{\|\cdot\|}$  est un espace de Hilbert de tenseur.

Remarquons que  $\mathcal{R}_1(V)$  est dense dans  $V_{\|\cdot\|}$ , et pour tout  $r \geq (1, \dots, 1)$  on a  $\mathcal{R}_1(V) \subset \mathcal{T}_r(V)$ , alors  $\text{span} \mathcal{T}_r(V)$  est aussi dense dans  $V_{\|\cdot\|}$ .

**Définition 1.7.3.** (Crossnorm)

On appelle "crossnorm" toute norme  $\|\cdot\|$  sur  ${}_a \bigotimes_{j=1}^d V_j$  vérifiant :

$$\left\| \bigotimes_{j=1}^d v^j \right\| = \prod_{j=1}^d \|v^j\|_j, \text{ pour toute } v^j \in V_j (1 \leq j \leq d). \quad (1.5)$$

On remarque facilement à partir de la définition de la crossnorm que l'application :

$$\bigotimes : \times_{j=1}^d (V_j, \|\cdot\|_j) \longrightarrow \left( {}_a\bigotimes_{j=1}^d V_j, \|\cdot\| \right) \\ (v^1, v^2, \dots, v^d) \longmapsto \bigotimes((v^1, v^2, \dots, v^d))$$

où  $\bigotimes((v^1, v^2, \dots, v^d)) = \bigotimes_{j=1}^d v^j$  est continue.

On note par  $\|\cdot\|^*$ , le dual de la crossnorm  $\|\cdot\|$ . Si  $\|\cdot\|$  est une crossnorm sur  ${}_a\bigotimes_{j=1}^d V_j$  alors  $\|\cdot\|^*$  est aussi une crossnorm sur  ${}_a\bigotimes_{j=1}^d V_j^*$ , c'est-à-dire :

$$\left\| \bigotimes_{j=1}^d \varphi^j \right\|^* = \prod_{j=1}^d \|\varphi^j\|_j^*, \quad \text{pour toute } \varphi^j \in V_j^* (1 \leq j \leq d). \quad (1.6)$$

**Définition 1.7.4.** Soient  $(V)_j$  l'ensemble des espaces des Banach pour tout  $1 \leq j \leq d$ . On définit la norme  $\|\cdot\|_V$  sur  $V = {}_a\bigotimes_{j=1}^d V_j$  telle que :

$$\|v\|_V := \sup \left\{ \frac{|(\varphi^1 \otimes \dots \otimes \varphi^d)(v)|}{\prod_{j=1}^d \|\varphi^j\|_j^*} : 0 \neq \varphi^j \in V_j^*, 1 \leq j \leq d \right\}$$

pour tout  $v \in V$ .

Soit  $(v_n)_{n \in \mathbb{N}}$  une suite dans  $V$ . On dit que  $v_n$  converge faiblement vers  $v$  dans  $V$ , si :

$$\lim_{n \rightarrow \infty} \langle v_n, \varphi \rangle = \langle v, \varphi \rangle \quad \text{pour tout } \varphi \in V^*,$$

et on note dans ce cas,  $v_n \rightharpoonup v$ .

Soit  $M \subset V$  un sous-espace de  $V$ . On dit que  $M$  est faiblement fermé si pour tout suite  $v_n \in M$ , telle que  $v_n \rightharpoonup v$ , on a  $v \in M$ .

**Proposition 1.7.5.** Soit  $V_{\|\cdot\|}$  un espace de Banach de tenseur muni de la norme  $\|\cdot\| \geq \|\cdot\|_V$  sur  $V$ .

L'ensemble  $\mathcal{T}_r(V)$  est faiblement fermé.

Soit  $\Omega = \Omega_1 \times \dots \times \Omega_d \subset \mathbb{R}^d$ , et on note par  $\alpha \in \mathbb{N}^d$  le multi-indice tel que  $|\alpha| = \sum_{k=1}^d \alpha_k$ .  $D^\alpha = \partial_{x_1}^{\alpha_1} \dots \partial_{x_d}^{\alpha_d}(u)$  est la dérivée d'ordre  $|\alpha|$  de  $u(x_1, \dots, x_d)$ .

On introduit l'espace de Sobolev  $W^{m,p}(\Omega)$  pour  $1 \leq p < \infty$  tel que :

$$W^{m,p}(\Omega) = \{u \in L^p(\Omega) : D^\alpha(u) \in L^p(\Omega), 0 \leq |\alpha| \leq m\}$$

muni de la norme :

$$\|v\|_{m,p} = \sum_{0 \leq |\alpha| \leq m} \|D^\alpha(v)\|_{L^p(\Omega)}.$$

Dans le cas où on suppose  $V_k = W^{m,p}(\Omega_k)$ , muni du produit scalaire  $\|\cdot\|_{m,p;k}$  qui est défini par :

$$\|w\|_{m,p;k} = \sum_{j=0}^m \|\partial_{x_k}^j(w)\|_{L^p(\Omega_k)},$$

on a alors :

$$W^{m,p}(\Omega) =_{\|\cdot\|_{m,p}} \bigotimes_{j=1}^d W^{m,p}(\Omega_j)$$

et

$$W_0^{m,p}(\Omega) =_{\|\cdot\|_{m,p}} \bigotimes_{j=1}^d W_0^{m,p}(\Omega_j).$$

**Lemme 1.7.6.** *Pour  $1 < p < +\infty$  et  $\Omega = \Omega_1 \times \cdots \times \Omega_d$ , l'espace  $\mathcal{R}_1 \left( \bigotimes_{j=1}^d L^p(\Omega_j) \right)$  est faiblement fermé dans  $L^p(\Omega)$ .*

**Preuve.** Soit  $\{v_n\}_{n \in \mathbb{N}}$  une suite dans  $\mathcal{R}_1 \left( \bigotimes_{j=1}^d L^p(\Omega_j) \right)$  telle que  $v_n = \bigotimes_{j=1}^d v_n^{(j)}$ .

Supposons que  $v_n \rightharpoonup v \in L^p(\Omega)$  et montrons que  $v \in \mathcal{R}_1 \left( \bigotimes_{j=1}^d L^p(\Omega_j) \right)$ . Puisque  $v_n \rightharpoonup v \in L^p(\Omega)$  alors  $\{v_n\}_{n \in \mathbb{N}}$  est bornée dans  $L^p(\Omega)$ . Posons :

$$\tilde{v}_n^{(j)} = \begin{cases} \frac{v_n^{(j)}}{\|v_n^{(j)}\|_{L^p(\Omega_j)}} & \text{si } \|v_n^{(j)}\|_{L^p(\Omega_j)} \neq 0 \\ 0 & \text{sinon} \end{cases}$$

pour  $j \in \{1, \dots, d-1\}$ , et :

$$\tilde{v}_n^{(d)} = \begin{cases} \prod_{j=1}^{d-1} \|v_n^{(j)}\|_{L^p(\Omega_j)} v_n^{(d)} & \text{si toutes les } \|v_n^{(j)}\|_{L^p(\Omega_j)} \neq 0 \\ 0 & \text{sinon.} \end{cases}$$

Pour  $j \in \{1, \dots, d-1\}$ , toutes les  $\{\tilde{v}_n^{(j)}\}_{n \in \mathbb{N}}$  sont bornées dans  $L^p(\Omega_j)$ . D'autre part, on a :

$$\left\| \bigotimes_{j=1}^d \tilde{v}_n^{(j)} \right\|_{L^p(\Omega_j)} = \prod_{j=1}^d \|\tilde{v}_n^{(j)}\|_{L^p(\Omega_j)},$$

donc, ou bien  $\|\tilde{v}_n^{(d)}\|_{L^p(\Omega_d)} = 0$  ou

$$\|\tilde{v}_n^{(d)}\|_{L^p(\Omega_d)} = \frac{\left\| \bigotimes_{j=1}^d \tilde{v}_n^{(j)} \right\|_{L^p(\Omega)}}{\prod_{j=1}^{d-1} \|\tilde{v}_n^{(j)}\|_{L^p(\Omega_j)}} = \left\| \bigotimes_{j=1}^d \tilde{v}_n^{(j)} \right\|_{L^p(\Omega)} = \left\| \bigotimes_{j=1}^d v_n^{(j)} \right\|_{L^p(\Omega)},$$

c'est-à-dire que  $\left\{ \tilde{v}_n^{(d)} \right\}_{n \in \mathbb{N}}$  est aussi bornée dans  $L^p(\Omega_d)$ .

Donc on peut trouver une sous-suite d'indices  $(n_k)_k$  telle que pour tout  $j \in \{1, \dots, d\}$  on a :  $\tilde{v}_{n_k}^{(j)} \rightharpoonup v^{(j)}$  dans  $L^p(\Omega_j)$ , par suite  $\bigotimes_{j=1}^d \tilde{v}_{n_k}^{(j)} \rightharpoonup \bigotimes_{j=1}^d v^{(j)}$  dans  $\mathcal{D}'(\Omega)$ , c'est-à-dire :

$$\bigotimes_{j=1}^d v_{n_k}^{(j)} \rightharpoonup \bigotimes_{j=1}^d v^{(j)} \text{ dans } \mathcal{D}'(\Omega).$$

L'unicité de la limite permet de conclure que :  $v = \bigotimes_{j=1}^d v^{(j)} \in \mathcal{R}_1 \left( \bigotimes_{j=1}^d L^p(\Omega_j) \right)$ .

□

**Lemme 1.7.7.** Soient  $m \geq 0$ ,  $1 < p < +\infty$  et  $\Omega = \Omega_1 \times \dots \times \Omega_d$ . Pour tout  $k \in \{1, \dots, d\}$ , on a :

$$\bigotimes_{k=1}^d w_k \in W^{m,p}(\Omega) \iff w_k \in W^{m,p}(\Omega_k)$$

pour toute  $w_k : \Omega_k \rightarrow \mathbb{R}$  telle que  $\bigotimes_{k=1}^d w_k \neq 0$ .

**Preuve.** Supposons que  $w_k \in W^{m,p}(\Omega_k)$ , pour tout  $k \in \{1 \dots d\}$ , alors :

$$\begin{aligned} \|\bigotimes_{k=1}^d w_k\|_{m,p} &= \sum_{0 \leq |\alpha| \leq m} \prod_{k=1}^d \|\partial_{x_k}^{\alpha_k}(w_k)\|_{L^p(\Omega_k)} \\ &\leq \sum_{\alpha \in \{0, \dots, m\}^d} \prod_{k=1}^d \|\partial_{x_k}^{\alpha_k}(w_k)\|_{L^p(\Omega_k)} \\ &= \prod_{k=1}^d \left( \sum_{j=1}^m \|\partial_{x_k}^j(w_k)\|_{L^p(\Omega_k)} \right) \\ &= \prod_{k=1}^d \|w_k\|_{m,p;k} < +\infty, \end{aligned}$$

et par suite  $\otimes_{k=1}^d w_k \in W^{m,p}(\Omega)$ .

Inversement, supposons que  $\otimes_{k=1}^d w_k \in W^{m,p}(\Omega)$ , alors

$$\|\otimes_{k=1}^d w_k\|_{m,p} = \sum_{0 \leq |\alpha| \leq m} \|\mathcal{D}^\alpha \otimes_{k=1}^d (w_k)\|_{L^p(\Omega)} < +\infty,$$

alors,  $\|\mathcal{D}^\alpha \otimes_{k=1}^d (w_k)\|_{L^p(\Omega)} < +\infty, \forall 0 \leq |\alpha| \leq m$ .

D'une part, prenons  $\alpha = (0 \cdots 0)$ , nous obtenons :

$$\|\otimes_{k=1}^d w_k\|_{L^p(\Omega)} = \prod_{k=1}^d \|w_k\|_{L^p(\Omega_k)} < +\infty,$$

ce qui nous donne :  $\|w_k\|_{L^p(\Omega_k)} < +\infty, \forall k \in \{1 \cdots d\}$ .

D'autre part, prenons  $\alpha = (\cdots 0, j, 0 \cdots), \forall k \in \{1 \cdots d\}$ , nous obtenons :

$$\|\mathcal{D}^\alpha \otimes_{l=1}^d (w_k)\|_{L^p(\Omega)} = \|\partial_{x_k}^j (w_k)\|_{L^p(\Omega_k)} \prod_{l \neq k} \|w_l\|_{L^p(\Omega_l)} < +\infty,$$

donc,  $\|\partial_{x_k}^j (w_k)\|_{L^p(\Omega_k)} < +\infty, \forall j \in \{1 \cdots m\}$ , par suite  $w_k \in W^{m,p}(\Omega_k)$  pour tout  $k \in \{1 \cdots d\}$ .  $\square$

**Proposition 1.7.8.** *Avec les mêmes conditions que le lemme précédent, on a :*

$$\mathcal{R}_1 \left( \bigotimes_{j=1}^d W^{m,p}(\Omega_j) \right) = \left\{ \otimes_{k=1}^d w_k : w_k \in W^{m,p}(\Omega_k) \right\},$$

est faiblement fermé dans  $(W^{m,p}(\Omega), \|\cdot\|_{m,p})$ .

**Preuve.** On va traiter le cas où  $m \geq 1$ , parce que le cas où  $m = 0$  est résolu dans le lemme 1.7.6. Considérons une suite  $\{z_n\}_{n \in \mathbb{N}} \subset \mathcal{R}_1 \left( \bigotimes_{j=1}^d W^{m,p}(\Omega_j) \right)$  qui converge faiblement vers  $z \in W^{m,p}(\Omega)$ . Comme

$$\mathcal{R}_1 \left( \bigotimes_{j=1}^d W^{m,p}(\Omega_j) \right) \subset \mathcal{R}_1 \left( \bigotimes_{j=1}^d L^p(\Omega_j) \right)$$

et  $\mathcal{R}_1 \left( \bigotimes_{j=1}^d L^p(\Omega_j) \right)$  est faiblement fermé dans  $L^p(\Omega)$  d'après le lemme 1.7.6,

alors  $z \in \mathcal{R}_1 \left( \bigotimes_{j=1}^d L^p(\Omega_j) \right)$ . Ainsi, il existe  $w_k \in L^p(\Omega_k)$  pour  $k = 1, 2, \dots, d$  tels que  $z = \otimes_{k=1}^d w_k \in W^{m,p}(\Omega)$ . D'où  $w_k \in W^{m,p}(\Omega_k) \forall 1 \leq k \leq d$ , d'après le lemme 1.7.7 et en conclusion  $z = \otimes_{k=1}^d w_k \in \mathcal{R}_1 \left( \bigotimes_{j=1}^d W^{m,p}(\Omega_j) \right)$ .  $\square$

**Proposition 1.7.9.** Soient  $1 < p < \infty$ ,  $m \geq 0$  et  $\Omega = \Omega_1 \times \Omega_2 \cdots \Omega_d$ . L'espace  $\mathcal{R}_1 \left( {}_a \bigotimes_{j=1}^d W_0^{m,p}(\Omega_j) \right)$  est faiblement fermé dans  $L^p(\Omega)$ .

**Lemme 1.7.10.**  $\mathcal{R}_1 \left( {}_a \bigotimes_{j=1}^d W_0^{m,p}(\Omega_j) \right)$  est dense dans  $W_0^{1,p}(\Omega)$ .

## 1.8 PGD progressive

Les méthodes *PGD* consistent en la construction d'une suite  $(u_m)_{m \in \mathbb{N}} \subset V$  de la manière suivante :

- (i) Initialisation :  $u_0 := 0$ .
- (ii) Choisir  $\widehat{z}_m := \arg \min_{z \in \mathcal{R}_1(V)} J(u_{m-1} + z)$ .
- (iii) Progression

### Stratégie (c)

$$u_m := u_{m-1} + \widehat{z}_m.$$

### Stratégie (r)

- ★ Construire un sous-espace fermé de  $V_{\|\cdot\|}$  contenant  $\widehat{z}_m$ , noté  $U(\widehat{z}_m)$
- ★ Choisir  $\widetilde{z}_m := \arg \min_{z \in U(\widehat{z}_m)} J(u_{m-1} + z)$ .
- ★  $u_m := u_{m-1} + \widetilde{z}_m$ .

### Stratégie (ℓ)

- ★ Construire un sous-espace fermé de  $V_{\|\cdot\|}$  contenant  $u_{m-1} + \widehat{z}_m$ , noté  $U(u_{m-1} + \widehat{z}_m)$
- ★ Choisir  $u_m := \arg \min_{u \in U(u_{m-1} + \widehat{z}_m)} J(u)$ .

Ainsi, une suite  $\mathbf{u} := (u_m)_{m \in \mathbb{N}}$  de type PGD peut être représentée par une suite (peut-être finie) de symboles, notée par :

$$\widehat{\alpha}(\mathbf{u}) = \alpha_1 \alpha_2 \alpha_3 \cdots \alpha_m \cdots$$

où  $\alpha_m \in \{\ell, c, r\}$ , pour tout  $m = 1, 2, \dots$

### 1.8.1 L'opérateur p-laplacien

Soient  $\Omega$  un ouvert dans  $\mathbb{R}^d$  tel que  $\Omega = \Omega_1 \times \cdots \times \Omega_d$ , et  $V_{\|\cdot\|} = W_0^{1,p}(\Omega)$ , la fermeture de  $C_c^\infty(\Omega)$  dans  $W^{1,p}(\Omega)$  pour la norme de  $W^{1,p}(\Omega)$  avec  $p > 1$  et :

$$\|v\| = \|\nabla v\|_{L^p(\Omega)},$$

qui est équivalente à la norme  $\|\cdot\|_{1,p}$  sur  $W^{1,p}(\Omega)$  qui apparait dans le paragraphe précédent. On considère la forme linéaire continue  $-\Delta_p v$  définie pour toute  $\varphi \in W^{1,p}(\Omega)$  par :

$$\langle -\Delta_p v, \varphi \rangle := \int_{\Omega} |\nabla v|^{p-2} \nabla v \cdot \nabla \varphi dx,$$

où  $\Delta_p v = \operatorname{div}(|\nabla v|^{p-2} \nabla v)$  et on a  $\Delta_p v \in (W_0^{1,p}(\Omega))' = W^{-1,p'}(\Omega)$  où  $\frac{1}{p} + \frac{1}{p'} = 1$ . Avec ces notation, on peut introduire la définition suivante :

**Définition 1.8.1.** Soit  $f \in W^{-1,p'}(\Omega)$ . On dit que  $v$  est une solution faible de l'équation :

$$\begin{cases} -\Delta_p v &= f \text{ dans } \Omega, \\ v &= 0 \text{ sur } \Omega, \end{cases}$$

dans  $W_0^{1,p}(\Omega)$  si et seulement si :

$$\int_{\Omega} |\nabla v|^{p-2} \nabla v \cdot \nabla \varphi dx = \langle -\Delta_p v, \varphi \rangle, \quad \forall \varphi \in W_0^{1,p}(\Omega).$$

Il est clair que les solutions faibles du système sont les points critiques de la fonctionnelle  $J$  définie par :

$$\begin{aligned} J: W_0^{1,p}(\Omega) &\longrightarrow \mathbb{R} \\ u &\longmapsto J(u) = \int_{\Omega} |\nabla v|^p dx - \int_{\Omega} f v dx. \end{aligned}$$

**Proposition 1.8.2.** La fonctionnelle  $J$  est dérivable au sens de Fréchet et  $J'(v) = -\Delta_p v - f$ .

**Preuve:**

En effet, soit  $\varphi \in W_0^{1,p}(\Omega)$  fixé et soit la fonction

$$\begin{aligned} F: \mathbb{R} &\longrightarrow \mathbb{R} \\ t &\longmapsto F(t) = J(v + t\varphi), \end{aligned}$$



On a

$$\begin{aligned}
 F(t) &= \frac{1}{p} \int_{\Omega} |\nabla v + t \nabla \varphi|^p dx - \int_{\Omega} f(v + t \varphi) dx \\
 &= \frac{1}{p} \langle \nabla v + t \nabla \varphi, \nabla v + t \nabla \varphi \rangle^{\frac{p}{2}} - \langle f, v + t \varphi \rangle \\
 &= \frac{1}{p} \langle \nabla v, \nabla v \rangle^{\frac{p}{2}} \left[ 1 + \frac{2 \langle \nabla v, \nabla \varphi \rangle}{\langle \nabla v, \nabla v \rangle} + o(t^2) \right]^{\frac{p}{2}} - \langle f, v \rangle - t \langle f, \varphi \rangle \\
 &= \frac{1}{p} \langle \nabla v, \nabla v \rangle^{\frac{p}{2}} \left[ 1 + \frac{p}{2} \frac{2 \langle \nabla v, \nabla \varphi \rangle}{\langle \nabla v, \nabla v \rangle} + o(t^2) \right] - \langle f, v \rangle - t \langle f, \varphi \rangle \\
 &= F(0) + t \left[ |\nabla v|^{2(\frac{p}{2}-1)} \langle \nabla v, \nabla \varphi \rangle - \langle f, \varphi \rangle + o(t) \right] \\
 &= F(0) + t \left[ \langle |\nabla v|^{p-2} \nabla v, \nabla \varphi \rangle - \langle f, \varphi \rangle + o(t) \right] \\
 &= F(0) + t \left[ \langle -\Delta_p v - f, \varphi \rangle + o(t) \right].
 \end{aligned}$$

Donc  $F'(0) = \langle J'(v), \varphi \rangle = \langle -\Delta_p v - f, \varphi \rangle$  et par suite  $J$  est Fréchet différentiable et :

$$J'(v) = -\Delta_p v - f.$$

■

Pour montrer l'ellipticité de  $J$ , notons que pour tous  $u$  et  $v$  dans  $W_0^{1,p}$ , on a

$$\begin{aligned}
 \langle J'(u) - J'(v), u - v \rangle_{W^{-1,p'}, W_0^{1,p}} &= \langle -(\Delta_p u - \Delta_p v), u - v \rangle_{W^{-1,p'}, W_0^{1,p}} \\
 &= \int_{\Omega} \langle |\nabla u|^{p-2} \nabla u - |\nabla v|^{p-2} \nabla v, \nabla u - \nabla v \rangle_{\mathbb{R}^d} dx.
 \end{aligned}$$

Ainsi, l'ellipticité de  $J$  découle du lemme suivant :

**Lemme 1.8.3.** *Il existe  $C_p > 0$  telle que pour tous  $x$  et  $y$  dans  $\mathbb{R}^d$ , on a*

$$\langle |x|^{p-2} x - |y|^{p-2} y, x - y \rangle_{\mathbb{R}^d} \geq \begin{cases} C_p |x - y|^p & \text{si } p \geq 2 \\ C_p \frac{|x - y|^2}{(|x| + |y|)^{2-p}} & \text{si } 1 < p < 2, \end{cases}$$

où  $|\cdot|$  désigne la norme euclidienne de  $\mathbb{R}^d$ .

**Preuve :** Sans perte de généralité, on peut supposer par homogénéité que  $|x| = 1$  et  $|y| \leq 1$ . En outre, quitte à faire un changement de base dans  $\mathbb{R}^d$ , on peut considérer  $x = (1, 0, \dots, 0)$  et  $y = (y_1, y_2, 0, \dots, 0)$  dans  $\mathbb{R}^d$  tels que  $\sqrt{y_1^2 + y_2^2} \leq 1$ .

**Cas 1 :**  $1 < p < 2$ . Il est clair que l'inégalité demandée est équivalente à

$$\left\{ \left( 1 - \frac{y_1}{(y_1^2 + y_2^2)^{(2-p)/2}} \right) (1 - y_1) + \frac{y_2^2}{(y_1^2 + y_2^2)^{(2-p)/2}} \right\} \frac{(1 - \sqrt{y_1^2 + y_2^2})^{2-p}}{(1 - y_1)^2 + y_2^2} \geq C.$$

Mais,

$$1 - \frac{y_1}{(y_1^2 + y_2^2)^{(2-p)/2}} \geq \begin{cases} 1 - \frac{y_1}{|y_1|^{2-p}} \geq (p-1)(1-y_1) & \text{si } 0 \leq y_1 \leq 1 \\ 1 - y_1 \geq (p-1)(1-y_1) & \text{si } y_1 \leq 0 \end{cases}$$

alors

$$(p-1) \{ (1-y_1)^2 + y_2^2 \} \frac{(1+y_1+y_2)^{(2-p)/2}}{(1-y_1)^2 + y_2^2} \geq p-1.$$

**Cas 2 :**  $p \geq 2$  Il suffit de prouver que

$$\frac{\left[ 1 - y_1 (y_1^2 + y_2^2)^{\frac{p-2}{2}} \right] (1-y_1) + y_2^2 (y_1^2 + y_2^2)^{\frac{p-2}{2}}}{[(1-y_1)^2 + y_2^2]^{\frac{p}{2}}} \geq C.$$

Notons  $t = \frac{|y|}{|x|}$  et  $s = \frac{\langle x, y \rangle}{|x||y|}$ , alors nous devons démontrer que la fonction :

$$f(t, s) = \frac{1 - (t^{p-1} + t) s + t^p}{(1 - 2ts + t^2)^{\frac{p}{2}}},$$

est bornée inférieurement. Pour  $t$  fixé, un calcul direct montre que :  $\frac{\partial f}{\partial s} = 0$  si

$$1 - (t^{p-1} + t) s + t^p = \frac{t^{p-2} + 1}{p} (1 - 2ts + t^2),$$

et pour ce point critique  $s$  de la fonction  $f$ , on a :

$$\begin{aligned} f(t, s) &= \frac{t^{p-2} + 1}{p} \frac{1}{(1 - 2ts + t^2)^{\frac{p-2}{2}}} \\ &\geq \frac{1}{p} \frac{t^{p-2} + 1}{(t+1)^{p-2}} \\ &\geq \frac{1}{p} \min_{0 \leq t \leq 1} \frac{t^{p-2} + 1}{(t+1)^{p-2}} \\ &\geq \frac{1}{2p}. \end{aligned}$$

Ce qui termine la preuve. □

## 1.9 Plan de la thèse

Cette thèse traite des méthodes de réduction de modèle. Elle est constituée de 4 chapitres qui peuvent être resumés comme suit :

### 1.9.1 Résumé des chapitres

#### **Chapitre 2 : La sensibilité des éléments propres d'opérateurs compacts auto-adjoints et quelques applications**

Dans ce chapitre, on va présenter des résultats portant sur la sensibilité des valeurs propres et des vecteurs propres d'opérateurs compacts auto-adjoints. On montre que les valeurs propres d'un opérateur compact auto-adjoint dépendent d'une manière *lipschitzienne* de la perturbation de l'opérateur, alors que les sous-espaces propres en dépendent de façon *localement lipschitzienne*.

Nos résultats présentent la sensibilité des espaces propres de dimension quelconque qui est une généralisation des travaux de Rousselet et Chenaïs [92] pour les espaces propres de dimension 1. Ainsi, ces résultats simplifient les travaux de Davis et Kahan [33] développés pour les perturbations des opérateurs hermitiens généraux.

D'autre part, la sensibilité de la base POD est mise en oeuvre dans le cas de perturbations temporelles, spatiales (Gappy-POD) ou bien paramétriques (variation de la viscosité, du nombre de Reynolds, etc ...).

#### **Chapitre 3 : PGD et Flot Optique**

Un modèle de calcul pour le flot optique a été introduit par Horn & Schunck [47] dans le cadre de l'espace de Sobolev  $H^1(\Omega)$ , où  $\Omega$  est un ouvert borné de  $\mathbb{R}^2$ .

Ce modèle ne profite pas du caractère tensoriel de l'espace  $H^1(\Omega)$ , lorsque  $\Omega$  est un produit cartésien d'intervalles réels. Dans ce travail, on montre comment utiliser ce caractère tensoriel de l'espace  $H^1(\Omega)$  et réduire de manière significative, le nombre d'inconnus (ordre  $N_x + N_y$  au lieu du nombre total de pixels  $N_x \times N_y$ ) et d'améliorer les propriétés de diffusion du mouvement dans les zones homogènes dont souffre la méthode de Horn & Schunck.

Nous montrons la convergence de l'algorithme PGD dans le cas du calcul du flot optique. Une étude numérique des deux approches, PGD et Horn&Schunck, est réalisée et la supériorité de la méthode PGD est mise en évidence aussi bien sur l'aspect temps de calcul que l'aspect robustesse de la méthode.

#### **Chapitre 4 : PGD sur la variété de Nehari pour les problèmes concave-convexe**

Dans ce chapitre, nous étudions la méthode PGD sur la variété de Nehari associée à une énergie  $E_\lambda$  définie sur un espace de Banach  $W$ . Cette variété est

définie par :

$$\mathcal{N} := \{u \in W \setminus \{0\} : E'_\lambda(u) \cdot u = 0\}.$$

Nous allons étudier un problème aux limites, non linéaire, concave-convexe, défini par :

$$\begin{cases} -\Delta_p u = \lambda |u|^{q-2}u + |u|^{r-2}u & \text{dans } \Omega, \\ u = 0, & \text{sur } \partial\Omega, \end{cases} \quad (1.7)$$

Notons que l'énergie associée à ce problème n'est ni convexe, ni coercive et ni bornée inférieurement. Nous démontrerons des résultats de convergence, pour la topologie faible, des suites PGD (lorsqu'elles sont bien définies) vers deux solutions extrémales sur la variété de Nehari. Plusieurs questions mathématiques concernant la PGD restent ouvertes dans ce chapitre. Ces questions font parties de nos perspectives de recherche.

Nous montrerons l'existence de deux branches de solutions, par rapport au paramètre  $\lambda$  pour ce problème en introduisant la contrainte naturelle  $\mathcal{N}$ . Ensuite, nous mettrons en évidence les liens de convergence des suites PGD associées aux valeurs d'adhérence de celles qui appartiennent à la variété de Nehari afin de conclure les résultats démontrés.

\* — \* — \* — \* — \* — \* — \*



# Chapitre 2

## La sensibilité des éléments propres d'opérateurs compacts auto-adjoints et quelques applications

### Sommaire

---

<b>2.1</b>	<b>Introduction</b>	<b>30</b>
<b>2.2</b>	<b>Préliminaires et notations</b>	<b>31</b>
<b>2.3</b>	<b>Régularité des valeurs propres et des vecteurs propres sur <math>\mathcal{K}_s(H)</math></b>	<b>32</b>
<b>2.4</b>	<b>Application à la méthode POD</b>	<b>37</b>
2.4.1	Variation du paramètre	38
2.4.2	La variation du domaine spatial : Gappy POD	40
2.4.3	La variation de l'intervalle temporel	41

---

## 2.1 Introduction

L'objectif principal de la POD est de déterminer une base *optimale* de dimension faible, pour représenter une donnée expérimentale de grande dimension ou toute simplement une donnée numérique.

Cette base de dimension faible permet de formuler des modèles d'ordre réduit pour les écoulements complexes. La POD permet de décomposer un champ d'écoulement spatio-temporel donné  $u(x, t)$  en un système orthogonal de modes spatiaux  $(\Phi_n(x))_{n \in \mathbb{N}}$  avec les amplitudes temporelles qui leur correspondent  $(a_n(t))_{n \in \mathbb{N}}$  comme suit :

$$u(x, t) = \sum_{n \geq 0} a_n(t) \Phi_n(x).$$

Le cadre mathématique classique de la POD est dans l'espace  $L^2([0, T]; H)$ , où  $[0, T]$  et  $H$  signifient respectivement un intervalle de temps et un espace de Hilbert des fonctions spatiales.

L'optimalité de la base  $(\Phi_n(x))_{n \in \mathbb{N}}$  signifie que toute expansion d'une série tronquée  $\sum_{n=0}^N a_n(t) \Phi_n(x)$  de  $u(x, t)$  dans cette base, a la plus petite erreur de troncature par le moyen quadratique temporel entre toutes les autres bases pour la norme  $L^2([0, T]; H)$ .

En utilisant la théorie spectrale pour les opérateurs compacts, auto-adjoints, la POD fournit un ordre naturel pour les modes spatiaux par rapport à la moyenne quadratique de leur amplitude.

Dans différents domaines de la mécanique de fluides ou Interaction Fluide Structure (IFS), la base POD est calculée pour un tel nombre de Reynolds ou bien pour un autre paramètre d'écoulement. Cette base est alors utilisée pour la construction d'un modèle réduit afin de chercher l'écoulement pour des autres valeurs de paramètres. Donc il est judicieux de se poser la question concernant le domaine de validité de ce modèle réduit. Ce domaine de validité a été précisé dans [1] *et al.* [39], pour les EDP paraboliques quasi-linéaires selon la régularité de la solution et le nombre de modes considérés. D'autres résultats similaires sont obtenus pour l'équation de Naviers-Stokes [1] et pour l'équation de Burgers [2]. Des études numériques sur la sensibilité paramétrique pour les problèmes de l'interaction de fluide structure ont été effectués dans [86]. Ces méthodes sont encore utilisées pour la théorie de contrôle [49, 56, 94], pour l'imagerie médicale [14, 91] où on cherche souvent à utiliser la base POD pour représenter des données malgré une variation, dans le domaine temporel, ou dans le domaine spatial, ou bien pour l'un des paramètres de l'équation étudiée.

Le but de ce chapitre est d'essayer à répondre à la problématique concernant la sensibilité de la base POD après une telle variation.

Dans les travaux de Davis et Kahane [33], les auteurs ont donné une réponse générale en ce qui concerne la rotation des espaces propres des opérateurs Hermitiens avec des perturbations Hermitiennes.

La théorie est développée dans [33], et a été simplifiée dans le cas des opérateurs compacts, auto-adjoints par Rousselet et Chenais [92]. Mais seulement la sensibilité des espaces propres de dimension 1, avec des variations compactes, a été étudiée. Dans ce travail, nous généralisons les résultats de [92] aux espaces propres multi-dimensionnels et nous mettons la sensibilité de la POD avec des variations temporelles, spatiales ou paramétriques, dans le cas général des perturbations d'un opérateur compact.

## 2.2 Préliminaires et notations

Soit  $H$  un espace de Hilbert, muni du produit scalaire  $\langle \cdot, \cdot \rangle_H$  et de la norme  $\| \cdot \|_H$ . Considérons  $\mathcal{K}_s(H)$  l'espace des opérateurs compacts auto-adjoints. On sait que pour tout  $K \in \mathcal{K}_s(H)$ , il existe une suite réelle des valeurs propres qui tend vers 0. De plus, 0 peut être une valeur propre si  $K$  n'est pas injectif.

Notons par  $(\mu_n(K))_{n \geq 1}$ , la suite des valeurs propres *non-nulles* de  $K$ , et  $m_n \in \mathbb{N}^*$ , la multiplicité de  $\mu_n(K)$ ; pour tout  $n \geq 1$ .

- $\mu_1^+, \dots, \mu_n^+, \dots$  est la suite décroissante des valeurs propres strictement positives :  $\mu_1^+ \geq \mu_2^+ \geq \dots \geq \mu_n^+ \geq \dots > 0$ .
- $\mu_1^-, \dots, \mu_n^-, \dots$  est la suite croissante des valeurs propres strictement négatives :  $\mu_1^- \leq \mu_2^- \leq \dots \leq \mu_n^- \leq \dots < 0$ .
- $E(\mu_n)$  est l'espace propre associé à  $\mu_n$ ,  $\dim E(\mu_n) = m_n$ .
- $E_{\text{nor}}(\mu_n) = E(\mu_n) \cap \mathbb{S}$ , où  $\mathbb{S}$  est la sphère unité de  $H$ .
- Pour tout  $u \in \mathbb{S}$ , posons  $d(u, E_{\text{nor}}(\mu_n)) = \min\{\|u - v\|_H : v \in E_{\text{nor}}(\mu_n)\}$ , la "distance" entre  $u$  et l'espace  $E_{\text{nor}}(\mu_n)$ .
- Pour tout sous-espace *fermé*  $E$  de  $H$ , notons par  $\prod_E$  la projection orthogonale sur  $E$ .

On dispose d'une caractérisation des valeurs propres grâce au théorème de Courant-Fisher :

$$\mu_n^+ = \min_{V \in \mathbb{H}_{n-1}} \max_{v \in V^\perp \cap \mathbb{S}} \langle K v, v \rangle_H, \quad (2.1)$$

où  $\mathbb{H}_{n-1}$  est la collection de tous les sous-espaces de  $H$  de dimension  $N - 1$ .

Cette valeur *min-max* est réalisée pour  $V = E(\mu_1^+) + E(\mu_2^+) + \dots + E(\mu_{n-1}^+)$  et pour tout vecteur  $v \in E_{\text{nor}}(\mu_n^+)$ . De la même manière,

$$\mu_n^- = \max_{V \in \mathbb{H}_{n-1}} \min_{v \in V^\perp \cap \mathbb{S}} \langle K v, v \rangle_H. \quad (2.2)$$



On note par  $V^\perp$  l'orthogonal, dans  $H$ , du sous-espace  $V$ . Cette valeur *max-min* est réalisée pour  $V = E(\mu_1^-) + E(\mu_2^-) + \cdots + E(\mu_{n-1}^-)$  et pour tout vecteur  $v \in E_{\text{nor}}(\mu_n^-)$ , et :

$$\|K\| = \max(\mu_1^+, -\mu_1^-) = \max \left( \max_{v \in \mathbb{S}} \langle K v, v \rangle_H, -\min_{v \in \mathbb{S}} \langle K v, v \rangle_H \right), \quad (2.3)$$

où  $\|\cdot\|$  représente la norme dans  $\mathcal{L}(H)$ , l'espace des opérateurs linéaires continus sur  $H$ .

## 2.3 Régularité des valeurs propres et des vecteurs propres sur $\mathcal{K}_s(H)$

Soit  $\mu_n$  la fonction définie pour tout  $n \in \mathbb{N}^*$ , par :

$$\begin{aligned} \mu_n : \mathcal{K}_s(H) &\longrightarrow \mathbb{R} \\ K &\longmapsto \mu_n(K) \end{aligned}$$

**Lemme 2.3.1.** *La fonction  $\mu_n$  est lipschitzienne de  $(\mathcal{K}_s(H), \|\cdot\|)$  dans  $\mathbb{R}$ .*

**Preuve.** Pour tout  $K_1, K_2$  dans  $\mathcal{K}_s(H)$  et  $u \in \mathbb{S}$ , l'inégalité :

$$|\langle K_2 u, u \rangle_H - \langle K_1 u, u \rangle_H| = |\langle (K_2 - K_1) u, u \rangle_H| \leq \|K_2 - K_1\|,$$

est satisfaite. Ce qui nous donne :

$$\langle K_1 u, u \rangle_H - \|K_2 - K_1\| \leq \langle K_2 u, u \rangle_H \leq \langle K_1 u, u \rangle_H + \|K_2 - K_1\|.$$

En appliquant la caractérisation min-max (2.1), on aura :

$$\mu_n^+(K_1) - \|K_2 - K_1\| \leq \mu_n^+(K_2) \leq \mu_n^+(K_1) + \|K_2 - K_1\|.$$

On peut utiliser les mêmes arguments pour les valeurs propres négatives  $\mu_n^-(K_i)$ ,  $i = 1, 2$ , avec la caractérisation de max-min (2.2), ce qui termine la démonstration.  $\square$

Dans ce qui suit, on va démontrer la régularité des vecteurs propres dans le cas des opérateurs compacts auto-adjoints. Avant d'énoncer notre résultat principal, on va introduire quelques notions de dépendance localement lipschitzienne pour les espaces propres dans  $\mathcal{K}_s(H)$ .

Soit  $K \in \mathcal{K}_s(H)$  et  $v \in \mathbb{S}$ . D'abord, on remarque que pour tout  $n \in \mathbb{N}^*$ , on a :

$$0 \leq d(v, E_{\text{nor}}(\mu_n(K))) \leq \sqrt{2} \quad (2.4)$$

où :

$$d(v; E_{\text{nor}}(\mu_n(K))) = \min \{ \|v - w\| ; w \in E_{\text{nor}}(\mu_n(K)) \}.$$

**Remarque 2.3.2.** *Le minimum est atteint, car la norme  $\|\cdot\|$  est une fonction continue et  $E(\mu_n(K)) \cap \mathbb{S}$  est compact.*

**Lemme 2.3.3.** *On a les équivalences suivantes :*

- i.  $d(v, E_{\text{nor}}(\mu_n(K))) = 0 \iff v \in E_{\text{nor}}(\mu_n(K)).$
- ii.  $0 < d(v, E_{\text{nor}}(\mu_n(K))) < \sqrt{2} \iff \hat{u} := \prod_{E(\mu_n(K))}(v) \neq 0$ , et  $u := \frac{\hat{u}}{\|\hat{u}\|}$  est l'unique point de  $E_{\text{nor}}(\mu_n(K))$  tel que  $d(v, E_{\text{nor}}(\mu_n(K))) = \|v - u\|.$
- iii.  $d(v, E_{\text{nor}}(\mu_n(K))) = \sqrt{2} \iff v \perp E_{\text{nor}}(\mu_n(K)).$

**Démonstration :**

- i. Si  $d(v, E_{\text{nor}}(\mu_n(K))) = 0$  alors  $\min \{\|v - w\| ; w \in E_{\text{nor}}(\mu_n(K))\}$ , et comme la norme est une fonction continue sur  $E_{\text{nor}}(\mu_n(K))$  qui est compact, alors le minimum est atteint par un  $u \in E_{\text{nor}}(\mu_n(K))$ .  
C'est-à-dire  $\|v - u\| = 0$  donc  $v = u \in E_{\text{nor}}(\mu_n(K)).$

Pour le deuxième sens, si  $v \in E_{\text{nor}}(\mu_n(K))$  alors :

$$0 = \|v - v\| = \min \{\|v - w\| ; w \in E_{\text{nor}}(\mu_n(K))\} = d(v, E_{\text{nor}}(\mu_n(K))).$$

- ii. Soient  $v \in \mathbb{S}$ , tel que  $0 < d(v, E_{\text{nor}}(\mu_n(K))) < \sqrt{2}$  et  $\hat{u} = \prod_{E(\mu_n(K))}(v)$ .  
Supposons que  $\hat{u} = 0$ , alors  $v \perp E(\mu_n(K))$ , alors quelque soit  $w$  dans  $E_{\text{nor}}(\mu_n(K))$ , on a  $\|v - w\| = \sqrt{2}$ , donc  $d(v, E_{\text{nor}}(\mu_n(K))) = \sqrt{2}$  absurde, donc  $\hat{u} \neq 0$ .  
Soit  $u = \frac{\hat{u}}{\|\hat{u}\|}$  et on va montrer que  $d(v, E_{\text{nor}}(\mu_n(K))) = \|v - u\|.$   
Quelque soit  $w \in E_{\text{nor}}(\mu_n(K))$ , on a :

$$\begin{aligned} \|v - w\|^2 &= \|v - \hat{u}\|^2 + \|w - \hat{u}\|^2 \\ &\geq \|v - \hat{u}\|^2 + \left\| \frac{\hat{u}}{\|\hat{u}\|} - \hat{u} \right\|^2 \\ &= \left\| v - \frac{\hat{u}}{\|\hat{u}\|} \right\|^2. \end{aligned}$$

Alors  $\frac{\hat{u}}{\|\hat{u}\|}$  réalise car elle est dans  $E_{\text{nor}}(\mu_n(K)).$

D'autre part, si  $\hat{u} := \prod_{E_{\text{nor}}(\mu_n(K))}(v) \neq 0$ , et si  $u := \frac{\hat{u}}{\|\hat{u}\|}$  est l'unique point de  $E_{\text{nor}}(\mu_n(K))$  tel que  $d(v, E_{\text{nor}}(\mu_n(K))) = \|v - u\|$  alors, on a :  
 $d(v, E_{\text{nor}}(\mu_n(K))) = \|v - u\| > 0$  car  $u \neq v$  et d'un autre côté, on a :

$$d(v, E_{\text{nor}}(\mu_n(K))) = \|v - u\| \leq (\|v\| + \|u\|)^{\frac{1}{2}} < \sqrt{2}.$$

- iii. Soit  $v \in \mathbb{S}$  tel que  $d(v, E_{\text{nor}}(\mu_n(K))) = \sqrt{2}$ . Supposons que  $v$  n'est pas perpendiculaire à  $E_{\text{nor}}(\mu_n(K))$ , d'après ii, on aura  $0 < d(v, E_{\text{nor}}(\mu_n(K))) < \sqrt{2}$ , absurde.

D'autre part, si  $v \perp E_{\text{nor}}(\mu_n(K))$ , alors  $\|v - w\|^2 = \|v\|^2 + \|w\|^2 = 1 + 1 = 2$  pour tout  $w \in E_{\text{nor}}(\mu_n(K))$ .

Il est clair que l'opérateur  $K - \mu_n(K) \text{id}_H$  n'est pas coercif sur  $H$ . Cependant, on peut montrer un résultat de coercivité de  $K - \mu_n(K) \text{id}_H$  sur  $E_{\text{nor}}(\mu_n(K))^\perp$ , pour tout  $K \in \mathcal{K}_s(H)$  et  $n \in \mathbb{N}^*$ .

**Lemme 2.3.4.** *Soit  $K \in \mathcal{K}_s(H)$  et  $n \in \mathbb{N}^*$ . Alors, il existe une constante  $C(K, n) > 0$  telle que :*

$$|\langle (K - \mu_n(K) \text{id}_H) w, w \rangle_H| \geq C(K, n) \|w\|^2, \text{ quelque soit } w \in E(\mu_n(K))^\perp. \quad (2.5)$$

**Preuve.** Considérons les sous-espaces :

- $E_n^0(K) = \text{Ker}(K)$ , le noyau de l'opérateur  $K$ .
- $E_n^+(K)$  est la somme de tous les espaces propres de  $K$  dont les valeurs propres sont plus grandes que  $\mu_n(K)$ .
- $E_n^-(K)$  est la somme de tous les espaces propres de  $K$  dont les valeurs propres sont plus petites que  $\mu_n(K)$ .

Puisque  $K$  est un opérateur compact, auto-adjoint, l'espace  $E(\mu_n(K))^\perp$  peut se décomposer de cette manière :

$$E(\mu_n(K))^\perp = E_n^0(K) \oplus E_n^+(K) \oplus E_n^-(K).$$

De plus, pour tout  $w \in E(\mu_n(K))^\perp$ ,

- si  $w \in E_n^0(K)$  alors  $\langle (K - \mu_n(K) \text{id}_H) w, w \rangle_H = -\mu_n(K) \|w\|^2$ ,
- si  $w \in E_n^+(K)$  alors  $\langle (K - \mu_n(K) \text{id}_H) w, w \rangle_H \geq (\overline{\mu}_n(K) - \mu_n(K)) \|w\|^2$ , où

$$\overline{\mu}_n := \min \{ \mu_k(K) : \mu_k(K) > \mu_n(K), k \in \mathbb{N}^* \},$$

- si  $w \in E_n^-(K)$  alors  $\langle (K - \mu_n(K) \text{id}_H) w, w \rangle_H \leq (\underline{\mu}_n(K) - \mu_n(K)) \|w\|^2$ , où

$$\underline{\mu}_n := \max \{ \mu_k(K) : \mu_k(K) < \mu_n(K), k \in \mathbb{N}^* \}.$$

La preuve se termine en posant :

$$C(K, n) := \min \{ |\mu_n(K)|, \overline{\mu}_n(K) - \mu_n(K), \mu_n(K) - \underline{\mu}_n(K) \} > 0. \quad (2.6)$$

□

**Remarque 2.3.5.** *La constante  $C(K, n)$  définie par (2.6) vérifie :*

$$\lim_{n \rightarrow +\infty} C(K, n) = 0, \quad (2.7)$$

*quelque soit  $K \in \mathcal{K}_s(H)$ .*

À ce stade, on peut montrer notre résultat principal.

Nous notons ici, que notre résultat est une généralisation d'un résultat obtenu par Rousselet et Chenais [92], où seulement la situation spéciale des vecteurs propres de dimension 1 est considérée. Notre travail fournit une démonstration simple et directe du résultat célèbre obtenu par Davis et Kahane [33] sur la rotation des vecteurs propres par perturbation, pour des opérateurs Hermitiens plus généraux.

**Théorème 2.3.6.** *Soit  $K_1 \in \mathcal{K}_s(H)$  et  $n \in \mathbb{N}^*$ . Alors il existe une constante de Lipschitz  $L(K_1, n) > 0$  telle que :*

$$d(v, E_{\text{nor}}(\mu_n(K_1))) \leq \frac{4}{C(K_1, n)} \|K_1 - K_2\|, \text{ pour tout } v \in E_{\text{nor}}(\mu_n(K_2)) \text{ et } K_2 \in \mathcal{K}_s(H),$$

et  $C(K_1, n)$  est la constante vérifiée par (2.6).

**Preuve.** Soit  $K_1, K_2$  dans  $\mathcal{K}_s(H)$  et  $n \in \mathbb{N}^*$  et soit  $v$  un vecteur quelconque de  $E_{\text{nor}}(\mu_n(K_2))$ . Grâce au lemme 2.3.3, on peut distinguer trois situations :

**Situation 1 :** Si  $d(v, E_{\text{nor}}(\mu_n(K_1))) = 0$ , alors  $v \in E_{\text{nor}}(\mu_n(K_1))$ , ce qui termine la démonstration.

**Situation 2 :** Si  $0 < d(v, E_{\text{nor}}(\mu_n(K_1))) < \sqrt{2}$ , en utilisant des arguments géométriques classiques, on peut démontrer facilement que,  $\hat{u} := \prod_{E(\mu_n(K_1))}(v) \neq 0$  et  $u := \frac{\hat{u}}{\|\hat{u}\|}$  est l'unique point dans  $E_{\text{nor}}(\mu_n(K_1))$  tel que  $d(v, E_{\text{nor}}(\mu_n(K_1))) = \|v - u\|$ . Donc :

$$v - u = (\|\hat{u}\| - 1)u + (v - \hat{u}), \text{ où } v - \hat{u} \perp E(\mu_n(K_1)). \quad (2.8)$$

Posons  $\alpha = \|\hat{u}\| - 1 = \langle v - u, u \rangle_H$ , alors, d'une part on a :

$$\|v - u\|^2 = \alpha^2 + \|v - \hat{u}\|^2. \quad (2.9)$$

et d'autre part on a  $v = u + (v - u)$ , donc on aura :

$$\|v\|^2 = \|u\|^2 + \|v - u\|^2 + 2\alpha,$$

avec :

$$\alpha = -\frac{1}{2}\|v - u\|^2.$$

En en substituant  $\alpha$  par sa valeur dans (2.9), on trouve :

$$\|v - u\|^2 \left(1 - \frac{1}{4}\|v - u\|^2\right) = \|v - \hat{u}\|^2.$$

Or on sait que  $\|v - u\| = d(v, E_{\text{nor}}(\mu_n(K_1))) \leq \sqrt{2}$ , c'est-à-dire  $(1 - \frac{1}{4}\|v - u\|^2) \geq \frac{1}{2}$ , donc :

$$\|v - u\|^2 \leq 2 \|v - \hat{u}\|^2.$$

En plus, l'inégalité de coercivité donne :

$$|\langle (K_1 - \mu_n(K_1) \text{id}_H)(v - \hat{u}), v - \hat{u} \rangle_H| \geq \frac{C(K_1, n)}{2} \|v - u\|^2.$$

Et l'égalité (2.8) permet de déduire :

$$\langle (K_1 - \mu_n(K_1) \text{id}_H)(v - \hat{u}), v - \hat{u} \rangle_H = \langle (K_1 - \mu_n(K_1) \text{id}_H)(v - u), v - u \rangle_H.$$

Pour simplifier les notations, posons  $\delta K := K_2 - K_1$  et  $\delta \mu_n = \mu_n(K_2) - \mu_n(K_1)$ , alors un calcul direct nous donne :

$$\langle (K_1 - \mu_n(K_1) \text{id}_H)(v - u), v - u \rangle_H = -\langle (\delta K - \delta \mu_n \text{id}_H)v, v - u \rangle_H,$$

et donc :

$$\begin{aligned} |\langle (K_1 - \mu_n(K_1) \text{id}_H)(v - u), v - u \rangle_H| &= |\langle (\delta K - \delta \mu_n \text{id}_H)v, v - u \rangle_H| \\ &\leq 2\|\delta K\| \times \|v\| \times \|v - u\| \\ &= 2\|\delta K\| \times \|v - u\| \end{aligned}$$

en utilisant la lemme 2.3.1. Enfin, on conclut que :

$$\frac{C(K_1, n)}{2} \|v - u\|^2 \leq 2\|K_2 - K_1\| \times \|v - u\|,$$

c'est-à-dire :

$$d(v, E_{\text{nor}}(\mu_n(K_1))) \leq \frac{4}{C(K_1, n)} \|K_2 - K_1\|.$$

**Situation 3 :** Si  $d(v, E_{\text{nor}}(\mu_n(K_1))) = \sqrt{2}$  alors  $v \in E_{\text{nor}}(\mu_n(K_1))^\perp$ . Des arguments similaires à ceux de la situation 2 avec  $\alpha = -1$  permettent de terminer la démonstration. □

Pour un tel  $K_1$ , la constante  $C(K_1, n)$  peut être plus petite que  $\|K_1 - K_2\|$ . Donc on peut trouver un vecteur  $v^* \in E_{\text{nor}}(\mu_n(K_2))$  tel que la distance  $d(v^*, E_{\text{nor}}(\mu_n(K_1)))$  est maximale, c'est-à-dire  $d(v^*, E_{\text{nor}}(\mu_n(K_1))) = \sqrt{2}$ . Dans de telles situations, l'espace propre  $E_{\text{nor}}(\mu_n(K_2))$  varie d'une façon significative même si la variation de l'opérateur  $\delta K := K_2 - K_1$  est petite.

Ici, un exemple illustre ce phénomène. Considérons la situation spatiale où  $H = \mathbb{R}^2$ ,

$$K_1 = \begin{pmatrix} 1 + \varepsilon & 0 \\ 0 & 1 - \varepsilon \end{pmatrix} \quad \text{et} \quad K_2 = \begin{pmatrix} 1 + \varepsilon & \varepsilon\sqrt{3} \\ \varepsilon\sqrt{3} & 1 - \varepsilon \end{pmatrix},$$

où  $\varepsilon$  est un petit paramètre réel. Les valeurs propres de  $K_1$  et  $K_2$  sont respectivement :

$$\mu_1(K_1) = 1 + \varepsilon, \mu_2(K_1) = 1 - \varepsilon, \mu_1(K_2) = 1 + 2\varepsilon, \text{ et } \mu_2(K_2) = 1 - 2\varepsilon,$$

dont la variation satisfait  $|\mu_i(K_2) - \mu_i(K_1)| = |\varepsilon| \ll 1$ , pour  $i = 1, 2$ . Cependant, les vecteurs propres de  $K_1$  et  $K_2$  sont respectivement :

$$\Phi_1(K_1) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \Phi_2(K_1) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad \Phi_1(K_2) = \begin{pmatrix} \sqrt{3}/2 \\ 1/2 \end{pmatrix} \text{ et } \Phi_2(K_2) = \begin{pmatrix} -\sqrt{3}/2 \\ 1/2 \end{pmatrix}$$

dont la variation satisfait  $\|\Phi_1(K_2) - \Phi_1(K_1)\| = \sqrt{2 - \sqrt{3}} \ll 1$  et  $\|\Phi_2(K_2) - \Phi_2(K_1)\| = 1 \ll 1$ .

Ceci nous conduit à introduire la valeur de *clustering* d'un opérateur  $K_1$  suite à une variation  $\delta K = \|K_2 - K_1\|$ .

**Définition 2.3.7.** *Considérons  $K_1, K_2$  dans  $\mathcal{K}_s(H)$  et  $\eta > 0$  un nombre réel positif.*

*La valeur d'index du clustering de l'opérateur  $K_1$ , pour la variation de l'opérateur  $\delta K$  et le seuil  $\eta > 0$ , est définie par :*

$$n_c(K_1, \delta K, \eta) := \min \left\{ n \in \mathbb{N}^* : \frac{\|\delta K\|}{C(K_1, n)} > \eta \right\}. \quad (2.10)$$

## 2.4 Application à la méthode POD

Une question significative pour la méthode POD est : étant donnée une solution d'un problème pour un ensemble de paramètres (physiques), est-ce que l'on peut contrôler la précision de la base POD correspondante lorsque l'espace de paramètres varie ? Dans ce qui suit, ceci sera notre sujet.

On va se limiter au cadre spatio-temporel, qui est constitué d'un domaine spatial  $\Omega \subset \mathbb{R}^N$ , d'un intervalle temporel  $[0, T]$  et d'un paramètre (physique)  $\lambda \in \Lambda$ , où  $\Lambda$  est un sous-espace d'un espace normé (ou plus généralement d'un espace métrique) donné. Soit  $H(\Omega)$ , un espace spatial de Hilbert, alors  $L^2(0, T; H(\Omega))$  est l'espace de toutes les fonctions  $u$  définies sur l'intervalle de temps  $[0, T]$  avec des valeurs dans  $H(\Omega)$ , telles que  $\int_0^T \|u(t)\|_{H(\Omega)}^2 dt < +\infty$ .

Considérons par exemple une famille de fonctions  $u \in L^2(0, T; H(\Omega))$ , dépendante du paramètre  $\lambda \in \Lambda$ . La famille  $(u_{\lambda, T, \Omega})_{\lambda, T, \Omega}$  représente la solution de l'EDP spatio-temporel, sur  $\Omega \times [0, T]$ , où  $H(\Omega)$  est l'espace des fonctions à "variable spatiale" et  $\lambda$  représente le paramètre physique.

La base POD associée à la fonction  $u_{\lambda, T, \Omega}$  est une base orthonormale  $(\Phi_n^{\lambda, T, \Omega})_{n \geq 1}$  de l'espace  $H(\Omega)$  qui maximise le moyen temporel de l'énergie cinétique contenue

dans la projection de  $u_{\lambda,T,\Omega}$  sur le sous-espace recouvert par  $(\Phi_n^{\lambda,T,\Omega})_{n \geq 1}$ . Notons que cette base est optimale au sens de :  $L^2(0, T; H(\Omega))$ .

Il est bien connu que  $\Phi_n^{\lambda,T,\Omega}$  est le vecteur propre associé à la  $n$ -ième valeur propre  $\mu_n$  de l'opérateur compact, auto-adjoint et positif, où :

$$\begin{aligned} K_{\lambda,T,\Omega} : H(\Omega) &\longrightarrow H(\Omega) \\ \varphi &\longmapsto \frac{1}{T} \int_0^T \langle u_{\lambda,T,\Omega}(t), \varphi \rangle_{H(\Omega)} u_{\lambda,T,\Omega}(t) dt, \end{aligned}$$

trié dans l'ordre décroissant  $\mu_1 \geq \mu_2 \geq \dots \geq \mu_n \geq \dots > 0$ . 0 peut être une valeur propre si  $K_{\lambda,T,\Omega}$  n'est pas injectif.

À ce stade là, on peut considérer trois types de variation : la variation du domaine spatial (Gappy-POD), la variation de l'intervalle de temps ou la variation du paramètre  $\lambda$ . Et la question est de savoir combien on peut étendre une base POD fixée pour représenter la solution dans le cas de ces variations. D'abord, on va décrire ces trois situations et les structures fonctionnelles correspondantes.

### 2.4.1 Variation du paramètre

Supposons que la solution  $u_{\lambda_1,T,\Omega} \in L^2(0, T; H(\Omega))$  est connue, pour une valeur de paramètre  $\lambda_1 \in \Lambda$ , où  $\Lambda$  est un espace normé et sa solution dépend d'une façon continue de  $\lambda$ , c'est-à-dire :

$$\lim_{\|h\|_{\Lambda} \rightarrow 0} \|u_{\lambda+h,T,\Omega} - u_{\lambda,T,\Omega}\|_{L^2(0,T;H(\Omega))} = 0.$$

Soit  $(\Phi_n^{\lambda_1,T,\Omega})_{n \geq 1}$ , une base POD associée à  $u_{\lambda_1,T,\Omega}$ , via l'opérateur compact, auto-adjoint et positif :

$$\begin{aligned} K_{\lambda_1,T,\Omega} : H(\Omega) &\longrightarrow H(\Omega) \\ \varphi &\longmapsto \frac{1}{T} \int_0^T \langle u_{\lambda_1,T,\Omega}(t), \varphi \rangle_{H(\Omega)} u_{\lambda_1,T,\Omega}(t) dt, \end{aligned}$$

et  $(\mu_n^{\lambda_1,T,\Omega})_{n \geq 1}$  respectivement ses valeurs propres triées dans l'ordre décroissant. On sait que :

$$\frac{1}{T} \left\| u_{\lambda_1,T,\Omega} - \sum_{n=1}^N \langle u_{\lambda_1,T,\Omega}, \Phi_n^{\lambda_1,T,\Omega} \rangle_{H(\Omega)} \Phi_n^{\lambda_1,T,\Omega} \right\|_{L^2(0,T;H(\Omega))}^2 = \sum_{n \geq N+1} \mu_n^{\lambda_1,T,\Omega}$$

et cette estimation est minimale parmi toutes les bases Hilbertiennes de  $H(\Omega)$ .

Maintenant, considérons une perturbation  $\delta\lambda$  de  $\lambda_1$  et posons  $\lambda_2 := \lambda_1 + \delta\lambda$ , la perturbation de l'opérateur résultant est ainsi  $\delta K = K_{\lambda_2, T, \Omega} - K_{\lambda_1, T, \Omega}$ . La dépendance continue de  $u_{\lambda_1, T, \Omega}$  par rapport à  $\lambda$  sur  $L^2(0, T; H(\Omega))$  conduit d'une façon directe à une dépendance continue de  $K_{\lambda, T, \Omega}$  par rapport à  $\lambda$  sur  $\mathcal{L}(H(\Omega))$  et donc :

$$\lim_{\|\delta\lambda\|_{\Lambda} \rightarrow 0} \|K_{\lambda_2, T, \Omega} - K_{\lambda_1, T, \Omega}\|_{\mathcal{L}(H(\Omega))} = 0.$$

Soient  $(\Phi_n^{\lambda_2, T, \Omega})_{n \geq 1}$  une base POD associée à  $u_{\lambda_2, T, \Omega}$ , via l'opérateur  $K_2$ . De la même façon, on a :

$$\frac{1}{T} \left\| u_{\lambda_2, T, \Omega} - \sum_{n=1}^N \langle u_{\lambda_2, T, \Omega}, \Phi_n^{\lambda_2, T, \Omega} \rangle_{H(\Omega)} \Phi_n^{\lambda_2, T, \Omega} \right\|_{L^2(0, T; H(\Omega))}^2 = \sum_{n \geq N+1} \mu_n^{\lambda_2, T, \Omega}.$$

On remarque un résultat étonnant lorsqu'on représente la solution  $u_{\lambda_2, T, \Omega}$  dans la première base  $(\Phi_n^{\lambda_1, T, \Omega})_{n \geq 1}$ . En effet, soit  $N > n_c(K_1, \delta K, \varepsilon)$ , pour un seuil fixe  $\varepsilon > 0$  et posons  $\beta_n := \langle u_{\lambda_2, T, \Omega}, \Phi_n^{\lambda_1, T, \Omega} \rangle_{H(\Omega)}$ , on obtient alors :

$$\begin{aligned} \left\| u_{\lambda_2, T, \Omega} - \sum_{n=1}^N \beta_n \Phi_n^{\lambda_1, T, \Omega} \right\|_{L^2(0, T; H(\Omega))}^2 &\leq \left\| u_{\lambda_2, T, \Omega} - \sum_{n=1}^{n_c(K_1, \delta K, \varepsilon)-1} \beta_n \Phi_n^{\lambda_1, T, \Omega} \right\|_{L^2(0, T; H(\Omega))}^2 \\ &= \left\| u_{\lambda_2, T, \Omega} - \sum_{n=1}^{n_c(K_1, \delta K, \varepsilon)-1} \beta_n \Phi_n^{\lambda_2, T, \Omega} \right\|_{L^2(0, T; H(\Omega))}^2 + O(\varepsilon), \\ &= \sum_{n \geq n_c(K_1, \delta K, \varepsilon)} \mu_n^{\lambda_2, T, \Omega} + O(\varepsilon), \end{aligned} \quad (2.11)$$

où  $O(\varepsilon)$  désigne une expression de l'ordre  $\varepsilon$ . La dernière égalité suggère que l'approximation de la solution  $u_{\lambda_2, T, \Omega}$ , dans le sous-espace de  $H(\Omega)$  engendré par les vecteurs propres  $(\Phi_n^{\lambda_1, T, \Omega})_{1 \leq n \leq N}$ , est limitée par l'index de clustering  $n_c(K_1, \delta K, \varepsilon)$ , même si sa dimension est plus grande que  $n_c(K_1, \delta K, \varepsilon)$ . Pour illustrer cela, notons par :

$$W_N^{\lambda_i, T, \Omega} := \text{span} \{ \Phi_n^{\lambda_i, T, \Omega} : 1 \leq n \leq N \}, \quad i = 1, 2$$

et

$$u_{\lambda_i, T, \Omega}^N := \prod_{W_N^{\lambda_i, T, \Omega}} u_{\lambda_i, T, \Omega}, \quad i = 1, 2,$$

la projection de  $u_{\lambda_i, T, \Omega}$  sur le sous-espace  $W_N^{\lambda_i, T, \Omega}$  par rapport au produit scalaire sur  $H(\Omega)$ , pour  $i = 1, 2$ .

Considérons la situation  $N > n_c(K_1, \delta K, \varepsilon)$ . En utilisant le théorème 2.3.6, on trouve :

$$\Phi_n^{\lambda_2, T, \Omega} = \Phi_n^{\lambda_1, T, \Omega} + O(\varepsilon), \quad \text{pour tout } n \in \{1, 2, \dots, n_c(K_1, \delta K, \varepsilon)\}.$$

Cependant, pour  $n = n_c(K_1, \delta K, \varepsilon)$ , on peut avoir  $\Phi_n^{\lambda_2, T, \Omega} = \Phi_n^{\lambda_1, T, \Omega} + O(1)$  [14]. Dans une telle situation, on aura :



$$\begin{aligned}
 \prod_{W_{n_c(K_1, \delta K, \varepsilon)}^{\lambda_1, T, \Omega}} u_{\lambda_2, T, \Omega} &= \sum_{n=1}^{n_c(K_1, \delta K, \varepsilon)} \langle u_{\lambda_2, T, \Omega}, \Phi_n^{\lambda_1, T, \Omega} \rangle_{H(\Omega)} \Phi_n^{\lambda_1, T, \Omega} \\
 &= \sum_{n=1}^{n_c(K_1, \delta K, \varepsilon)} \langle u_{\lambda_2, T, \Omega}, \Phi_n^{\lambda_2, T, \Omega} \rangle_{H(\Omega)} \Phi_n^{\lambda_2, T, \Omega} + O(\varepsilon) + \alpha,
 \end{aligned}$$

où :

$$\alpha = \langle u_{\lambda_2, T, \Omega}, \Phi_{n_c(K_1, \delta K, \varepsilon)}^{\lambda_1, T, \Omega} \rangle_{H(\Omega)} \Phi_{n_c(K_1, \delta K, \varepsilon)}^{\lambda_1, T, \Omega} - \langle u_{\lambda_2, T, \Omega}, \Phi_{n_c(K_1, \delta K, \varepsilon)}^{\lambda_2, T, \Omega} \rangle_{H(\Omega)} \Phi_{n_c(K_1, \delta K, \varepsilon)}^{\lambda_2, T, \Omega}.$$

D'où, on obtient :

$$\prod_{W_{n_c(K_1, \delta K, \varepsilon)}^{\lambda_1, T, \Omega}} u_{\lambda_2, T, \Omega} = \prod_{W_{n_c(K_1, \delta K, \varepsilon)}^{\lambda_2, T, \Omega}} u_{\lambda_2, T, \Omega} + O(\varepsilon) + \alpha.$$

Puisque  $\alpha$  n'a pas de raison d'être petite par rapport à  $\| \cdot \|_{H(\Omega)}$ , il s'ensuit que la distance entre la projection  $\prod_{W_{n_c(K_1, \delta K, \varepsilon)}^{\lambda_1, T, \Omega}} u_{\lambda_2, T, \Omega}$  et  $\prod_{W_{n_c(K_1, \delta K, \varepsilon)}^{\lambda_2, T, \Omega}} u_{\lambda_2, T, \Omega}$  peut être grande. Mais comme il était indiqué ci-dessus, on a :

$$\prod_{W_{n_c(K_1, \delta K, \varepsilon)-1}^{\lambda_1, T, \Omega}} u_{\lambda_2, T, \Omega} = \prod_{W_{n_c(K_1, \delta K, \varepsilon)-1}^{\lambda_2, T, \Omega}} u_{\lambda_2, T, \Omega} + O(\varepsilon),$$

ce qui signifie que lorsqu'on considère un sous-espace de vecteurs propres de la POD qui correspond aux valeurs propres clusturées, l'approximation de la solution est sensible pour des petites perturbations de paramètres. Dans ce qui suit, on va présenter une autre situation où la sensibilité de clustering peut exister. Plus précisément, on va voir que la méthode de Gappy POD ou, l'extension de l'intervalle de temps, vont perturber l'opérateur compact et donc, le phénomène des valeurs propres clusturées peut exister dans ce cas.

## 2.4.2 La variation du domaine spatial : Gappy POD

La méthode de gappy POD permet de reconstruire une donnée gappy (non complète) en utilisant une base POD. Cette méthode a été introduite par Everson et Sirovich [41] et on peut l'expliquer de la façon suivante : soit  $(\Phi_n^{\lambda, T, \Omega})_{n \geq 1}$  une base POD correspondante à une solution  $u_{\lambda, T, \Omega}$ . Soit  $\omega$  un "petit" sous-espace de  $\Omega$  et  $u_{\lambda, T, \Omega \setminus \omega}$ , la restriction de  $u_{\lambda, T, \Omega}$  à  $\Omega \setminus \omega$ . Le but de la méthode Gappy POD est de reconstruire  $u_{\lambda, T, \Omega \setminus \omega}$  sur le sous-ensemble manquant  $\omega$  en utilisant la base  $(\Phi_n^{\lambda, T, \Omega})_{n \geq 1}$  d'une façon directe. À cet effet, écrivons la solution "réparée" sur tout  $\Omega$  de la donnée non complète  $u_{\lambda, T, \Omega \setminus \omega}$  par :

$$\hat{u}_{\lambda, T, \Omega} := \sum_{n \geq 1} \hat{a}_n(t) \Phi_n^{\lambda, T, \Omega},$$

où les coefficients temporels inconnus  $(\widehat{a}_n(t))_{n \geq 1}$  minimisent le terme d'erreur  $\|\widehat{u}_{\lambda,T,\Omega} - u_{\lambda,T,\Omega \setminus \omega}\|_{H(\Omega)|_{\Omega \setminus \omega}}^2$  et,  $\|\cdot\|_{H(\Omega)|_{\Omega \setminus \omega}}^2$  signifie la restriction de la norme de  $H(\Omega)$  à  $\Omega \setminus \omega$ .

On va voir que la procédure de Gappy POD peut se placer dans le cas d'une perturbation de la base spatiale. En effet, soient  $K_1 := K_{\lambda,T,\Omega}$  et  $K_2 := K_{\lambda,T,\Omega \setminus \omega}$  les opérateurs compacts, auto-adjoints et positifs, associés à  $u_{\lambda,T,\Omega}$  et  $u_{\lambda,T,\Omega \setminus \omega}$  respectivement. Alors  $\delta K_\omega := K_1 - K_2$  est définie par :

$$\begin{aligned} \delta K_\omega : H(\Omega) &\longrightarrow H(\Omega) \\ \varphi &\longmapsto \frac{1}{T} \int_0^T \langle u_{\lambda,T,\omega}(t), \varphi \rangle_{H(\Omega)} u_{\lambda,T,\omega}(t) dt, \end{aligned}$$

où  $u_{\lambda,T,\omega}$  est la restriction de  $u_{\lambda,T,\Omega}$  à  $\omega$ . Il est clair que  $\delta K_\omega \in \mathcal{K}_s(H(\Omega))$  et maintenant, on peut trouver facilement avec le théorème de convergence dominée de Lebesgue :

$$\lim_{|\omega| \rightarrow 0} \delta K_\omega = 0 \quad \text{in } \mathcal{L}(H(\Omega)),$$

où  $|\omega|$  est la mesure de  $\omega$ .

### 2.4.3 La variation de l'intervalle temporel

Supposons que la solution  $u_{\lambda,T_1,\Omega} \in L^2(0, T_1; H(\Omega))$  est connue sur l'intervalle de temps  $[0, T_1]$ . Soit  $(\Phi_n^{\lambda_1, T_1, \Omega})_{n \geq 1}$ , la base POD associée à  $u_{\lambda_1, T_1, \Omega}$ , pour l'opérateur compact, auto-adjoint positif :

$$\begin{aligned} K_{\lambda, T_1, \Omega} : H(\Omega) &\longrightarrow H(\Omega) \\ \varphi &\longmapsto \frac{1}{T_1} \int_0^{T_1} \langle u_{\lambda, T_1, \Omega}(t), \varphi \rangle_{H(\Omega)} u_{\lambda, T_1, \Omega}(t) dt. \end{aligned}$$

Soit  $\delta T > 0$ , une perturbation de temps et fixons  $T_2 := T_1 + \delta T$ . La perturbation associée dans le terme de l'opérateur de corrélation est définie de cette façon :

$$\begin{aligned} \delta K_T : H(\Omega) &\longrightarrow H(\Omega) \\ \varphi &\longmapsto \int_0^{T_2} \langle u_{\lambda, T_2, \omega}(t), \varphi \rangle_{H(\Omega)} u_{\lambda, T_2, \omega}(t) \left( \frac{1}{T_2} - \frac{1}{T_1} \mathbf{1}_{[0, T_1]} \right) dt, \end{aligned}$$

où  $\mathbf{1}_{[0, T_1]}$  désigne la fonction indicatrice du sous-intervalle  $[0, T_1]$  dans  $[0, T_2]$ . Comme précédemment, l'opérateur  $\delta K_T \in \mathcal{K}_s(H(\Omega))$  et  $\lim_{\delta T \rightarrow 0} \delta K_T = 0$  dans  $\mathcal{L}(H(\Omega))$ .

Ainsi, la perturbation de l'intervalle de temps produit une perturbation de l'opérateur compact et donc une perturbation de la base POD. D'où, les résultats

de sensibilité similaires sont prévus dans le cas où les valeurs propres de l'opérateur  $K_{\lambda, T_1, \Omega}$  sont closturées.

$$* - * - * - * - * - * - * - * - *$$

# Chapitre 3

## PGD et Flot optique

### Sommaire

---

<b>Présentation du chapitre</b> . . . . .	<b>44</b>
<b>3.1 Introduction</b> . . . . .	<b>44</b>
<b>3.2 Matériels et méthodes</b> . . . . .	<b>45</b>
3.2.1 Approche PGD proposée pour les méthodes variationnelles . . . . .	45
3.2.2 Implementation des approches PGD et Horn&Schunk en flot optique . . . . .	54
3.2.3 Simulations numériques . . . . .	58
3.2.4 Evaluation de la méthode PGD proposée . . . . .	59
<b>3.3 Résultats</b> . . . . .	<b>59</b>
<b>3.4 Discussion</b> . . . . .	<b>61</b>
<b>3.5 Conclusion</b> . . . . .	<b>62</b>

---

### 3.1 Introduction

Les problèmes variationnels occupent une place importante en mathématiques depuis les problèmes isopérimétriques (philosophes grecs), les courbes optimales (L. Euler), la distribution optimale de charge électrique sur une surface conductrice (Principe de Dirichlet), les applications conformes (B. Riemann), les systèmes hamiltoniens (W. Hamilton), etc ...

Les travaux de D. Hilbert au début du 20-ème siècle ont donné la première justification rigoureuse du principe de Dirichlet et, la théorie moderne du calcul variationnel a connu un énorme développement depuis.

Bien que les théorèmes d'existence de solutions dans les problèmes de minimisation dans des espaces fonctionnels sophistiqués exigent des outils complexes d'analyse fonctionnelle, le principe peut être simplement énoncé : *" In any event therefore the integral will be nonnegative and hence there **must** exist a distribution (of electrical charge) for which this integral assumes its minimum value "* (Gauss - 1867).

Dans ce travail, on s'intéresse au calcul du flot optique dans un cadre classique d'espaces de Sobolev. Il s'agit d'introduire une énergie formée d'un terme traduisant *la contrainte de flot optique*, en plus, d'un terme de régularisation quadratique portant sur le gradient de ce flot. Ce modèle a été introduit et étudié par Horn & Schunck [47] dans le cadre de l'espace de Sobolev  $H^1(\Omega)$ , où  $\Omega$  est un ouvert borné de  $\mathbb{R}^2$ .

La méthode de Horn & Schunck ne profite pas du caractère tensoriel de l'espace  $H^1(\Omega)$ , lorsque  $\Omega$  est un produit cartésien d'intervalles réels. Dans le présent travail, on montre comment utiliser ce caractère tensoriel de l'espace  $H^1(\Omega)$  et réduire de manière significative, le nombre d'inconnus (ordre  $N_x + N_y$  au lieu du nombre total de pixels  $N_x \times N_y$ ) et d'améliorer même les propriétés de diffusion du flot dans les zones homogènes dont souffre la méthode de Horn & Schunck.

Bien entendu, ce travail peut être facilement étendu à d'autres thématiques qui intéressent la communauté de l'analyse d'image, tels les problèmes de segmentation/restauration d'image, acquisition/compression d'image, etc ...

En effet, en restauration d'image, on peut introduire une énergie à minimiser, formée d'un terme d'attache à l'observation (ou à la mesure) et d'un terme de régularisation pénalisant les solutions ne satisfaisant pas les critères de régularité désirés. De même, pour les problèmes de segmentation d'image, Mumford & Shah [74] ont proposé un modèle de segmentation où le problème de minimisation porte à la fois sur la recherche d'une fonction régulière par morceaux (représentant les objets), ainsi qu'un sous ensemble  $\Gamma \subset \Omega$  représentant l'ensemble des discontinuités. Plus précisément,

$$J_{MS}(u, \Gamma) = \int_{\Omega} (u - I)^2 dx + \alpha \int_{\Omega \setminus \Gamma} |\nabla u|^2 + \beta \int_{\Gamma} d\sigma.$$

Le premier terme de  $J_{MS}(u, \Gamma)$  contraint  $u$  à être proche de l'image de référence

*I.* Le second terme force le caractère lisse par morceaux de  $u$  : les variations de  $u$  sont pénalisées en dehors de l'ensemble de discontinuité  $\Gamma$ . Enfin le troisième terme est une sorte de régularisation des contours des objets (supposés à bords réguliers).

Nous nous sommes restreints à une approche quadratique en gradient des déplacements (pour avoir un cadre fonctionnel  $H^1(\Omega)$ ) pour bénéficier du caractère tensoriel de cet espace de Sobolev. Depuis le travail de Rudin-Osher-Fatemi [90], les approches variationnelles à croissance linéaires (à l'infini) en gradient, dont le cadre fonctionnel adéquat est  $BV(\Omega)$ , qui a de bonnes propriétés topologiques pour sa topologie faible  $*$ , contrairement à l'espace de Sobolev  $W^{1,1}(\Omega)$ , ont connu un vaste développement. Malheureusement, l'espace  $BV(\Omega)$  n'a pas une structure tensorielle, en effet, on peut montrer facilement que la fonction indicatrice du triangle  $ABC$  dans un carré  $\Omega \equiv ABCD$  ne peut être la limite d'une série tensorielle pour la norme  $BV(\Omega)$ . Pour ces raisons, d'un point de vue théorique, la méthode PGD présente de grandes difficultés pour établir des résultats de convergence, même si on arrive à bien définir la suite PGD, bien que les simulations numériques de telles approches puissent donner d'excellents résultats au niveau de la précision des estimations du flot optique et bien sûr au niveau temps de calcul également.

## 3.2 Matériels et méthodes

### 3.2.1 Approche PGD proposée pour les méthodes variationnelles

Cette section traite des problèmes de minimisation de fonctionnelles convexes définies sur des espaces de Banach tensoriels. La structure tensorielle de l'espace de Banach permet d'effectuer l'optimisation par des méthodes de réduction dimensionnelle de manière progressive. Dans ce qui suit, on s'intéresse aux méthodes de *décomposition progressive généralisée* (ou PGD) introduite par Ladevèze [62], puis Ammar *et al* dans [10, 11]. Plus précisément, on s'intéresse à l'étude mathématique de telles méthodes dans le cas variationnel et leur convergence.

#### Espaces de Banach tensoriels et PGD

Considérons une famille finie d'espaces de Banach réels  $(V_k, \|\cdot\|_k)_{1 \leq k \leq d}$ , où  $d \geq 2$  est un entier. L'espace tensoriel algébrique engendré par la famille  $(V_k, \|\cdot\|_k)_{1 \leq k \leq d}$ , noté

$$V := \bigotimes_{k=1}^d V_k,$$

est l'ensemble de toutes les combinaisons linéaires finies des tenseurs élémentaires  $v = \bigotimes_{k=1}^d v_k$ , avec  $v_k \in V_k$ . Le suffixe "a" dans " ${}_a \bigotimes$ " réfère à la nature "algébrique" de ce produit tensoriel.

On dit que  $V_{\|\cdot\|}$  est un espace de Banach tensoriel s'il existe un espace tensoriel algébrique  $V$  et une norme  $\|\cdot\|$  sur  $V$  tels que  $V_{\|\cdot\|}$  est le complété de  $V$  relativement à la norme  $\|\cdot\|$ , c'est à dire

$$V_{\|\cdot\|} = \overline{{}_a \bigotimes_{k=1}^d V_k}^{\|\cdot\|}.$$

Lorsque  $V_{\|\cdot\|}$  est un espace de Hilbert, on dira de manière similaire que  $V_{\|\cdot\|}$  est un espace de Hilbert tensoriel.

Dans la suite, l'ensemble des tenseurs élémentaires d'un espace tensoriel algébrique  $V$  sera noté par  $\mathcal{R}_1(V)$ . Ainsi l'espace engendré par  $\mathcal{R}_1(V)$  est  $V$  et par conséquent  $\mathcal{R}_1(V)$  est dense dans  $V_{\|\cdot\|}$ .

Considérons maintenant une fonctionnelle  $J : V_{\|\cdot\|} \rightarrow \mathbb{R}$ , où  $V_{\|\cdot\|}$  est un espace de Banach tensoriel *réflexif*. On s'intéresse au problème de minimisation

$$\inf_{u \in V_{\|\cdot\|}} J(u), \quad (3.1)$$

On fera les hypothèses suivantes :

- ( $H_1$ )  $J$  est différentiable et strictement convexe.
- ( $H_2$ )  $J$  est coercive.

En utilisant les arguments classiques du calcul variationnel, on montre que le problème (3.1) admet une unique solution, sous les hypothèses ( $H_1$ ) – ( $H_2$ ). Pour définir les méthodes PGD, on fait l'hypothèse complémentaire :

$$(\mathcal{H}_3) \quad \mathcal{R}_1(V) \text{ est faiblement fermé dans } V_{\|\cdot\|},$$

où  $V := {}_a \bigotimes_{k=1}^d V_k$  est l'espace tensoriel algébrique associé à  $V_{\|\cdot\|}$ .

Notons que pour tout  $u \in V$ , le problème

$$\inf_{z \in \mathcal{R}_1(V)} J(u + z) \quad (3.2)$$

possède au moins une solution, sous les hypothèses ( $H_1$ ) – ( $H_3$ ). L'éventuelle non-unicité de solution du problème (3.2) est due au fait que  $\mathcal{R}_1(V)$  n'est pas convexe.

Les méthodes *PGD* associées au problème (3.1) consistent en la construction d'une suite  $(u_m)_{m \in \mathbb{N}} \subset V$  de la manière suivante :

- (i) Initialisation :  $u_0 := 0$ .
- (ii) Choisir  $\widehat{z}_m := \arg \min_{z \in \mathcal{R}_1(V)} J(u_{m-1} + z)$ .
- (iii) Progression
  - Stratégie (c)
    - $u_m := u_{m-1} + \widehat{z}_m$ .
  - Stratégie (r)
    - ★ Construire un sous-espace fermé de  $V_{\|\cdot\|}$  contenant  $\widehat{z}_m$ , noté  $U(\widehat{z}_m)$
    - ★ choisir  $\widetilde{z}_m := \arg \min_{z \in U(\widehat{z}_m)} J(u_{m-1} + z)$ .
    - ★  $u_m := u_{m-1} + \widetilde{z}_m$ .
  - Stratégie (ℓ)
    - ★ Construire un sous-espace fermé de  $V_{\|\cdot\|}$  contenant  $u_{m-1} + \widehat{z}_m$ , noté  $U(u_{m-1} + \widehat{z}_m)$
    - ★ Choisir  $u_m := \arg \min_{u \in U(u_{m-1} + \widehat{z}_m)} J(u)$ .

Ainsi, une suite  $\mathbf{u} := (u_m)_{m \in \mathbb{R}}$  de type PGD peut être représentée par une suite (peut-être finie) de symboles, notée par :

$$\widehat{\alpha}(\mathbf{u}) = \alpha_1 \alpha_2 \alpha_3 \cdots \alpha_m \cdots$$

où  $\alpha_m \in \{\ell, c, r\}$ , pour tout  $m = 1, 2, \dots$

### Exemple on $L^2 - L^2$ Optical Flow

Soit  $\Omega$  un ouvert de  $\mathbb{R}^3$  qui s'écrit comme produit cartésien de trois intervalles ouverts  $(\Omega_k)_{1 \leq k \leq 3}$  de  $\mathbb{R}$ .

On notera la fonction "niveau de gris" au point  $(x, y, z) \in \Omega$  et à l'instant  $t$  par  $I(x, y, z, t)$  et

$$\Gamma(x_0, y_0, z_0, t_0) := \{(x(t), y(t), z(t)) : (x(t_0), y(t_0), z(t_0)) = (x_0, y_0, z_0), t \geq t_0\}$$

la trajectoire, supposée régulière, du point  $(x, y, z)$ . Ainsi, il est raisonnable de supposer que le niveau de gris reste "constant" sur un intervalle de longueur "infinitement petite" inclus dans  $\Gamma(x_0, y_0, z_0, t_0)$ . Cette hypothèse s'écrit mathématiquement :

$$\frac{dI}{dt}(x(t), y(t), z(t), t) = 0, \quad \forall t > t_0,$$

et elle s'appelle la contrainte de flot optique. Cette contrainte peut être réécrite sous la forme :

$$\sigma(x(t), y(t), z(t), t) \cdot \nabla I(x(t), y(t), z(t), t) + \frac{\partial I}{\partial t}(x(t), y(t), z(t), t) = 0, \quad \forall t > t_0,$$

où  $\sigma(x(t), y(t), z(t), t)$  est la vitesse du point  $(x(t), y(t), z(t))$  à l'instant  $t$  et  $\nabla I(x(t), y(t), z(t), t)$  est le gradient spatial de  $I$  au point  $(x(t), y(t), z(t))$ . Nous



noterons dans toute la suite

$$\sigma := (u, v, w).$$

On se fixe un instant  $t > t_0$ . Pour alléger les notations, on ne notera plus la dépendance en  $t$  des fonctions  $\sigma$  et  $I$ . Introduisons l'énergie

$$J_0(\sigma) := \int_{\Omega} \left( \sigma(x, y, z) \cdot \nabla I(x, y, z) + \frac{\partial I}{\partial t}(x, y, z) \right)^2 dx dy dz.$$

Il est classique d'introduire l'énergie suivante pour calculer une valeur approchée du flot optique :

$$J(\sigma) := J_0(\sigma) + J_1(\sigma), \quad (3.3)$$

où

$$J_1(\sigma) := \int_{\Omega} c(x, y, z) |\sigma(x, y, z)|^2 dx dy dz + \lambda \int_{\Omega} |\nabla \sigma(x, y, z)|^2 dx dy dz.$$

Le paramètre réel  $\lambda > 0$  est un paramètre de régularisation concernant  $\nabla \sigma$ , quant à la fonction  $c(x, y, z) \geq \varepsilon > 0$ , elle pénalise les zones homogènes, caractérisées par un faible gradient de l'image. Ainsi,  $c(x, y, z)$  est petite (resp. grande) dans les zones de  $\Omega$  où  $|\nabla I(x, y, z)|$  est grand (resp. petit). On supposera dans la suite que  $c \in \mathcal{C}^\infty(\Omega)$ .

Il est clair que le cadre fonctionnel pour la fonction  $J$  est  $W_{\|\cdot\|} := [H^1(\Omega)]^3$  et le problème de minimisation associé au calcul du flot optique est par conséquent :

$$\inf_{\sigma \in W_{\|\cdot\|}} J(\sigma). \quad (3.4)$$

Dans toute la suite, les normes usuelles sur  $W_{\|\cdot\|}$  et sur  $H^1(\Omega)$  seront notées par  $\|\cdot\|$ .

Ainsi on a

$$\|\xi\| = \begin{cases} \|\xi\|_2 + \|\nabla \xi\|_2 & \text{si } \xi \in H^1(\Omega) \\ \|\xi\|_2^{1/2} + \|\nabla \xi\|_2^{1/2} & \text{si } \xi \in W_{\|\cdot\|} \end{cases}$$

où  $\|\cdot\|_2$  désigne la norme usuelle de  $L^2(\Omega)$ ,  $|\cdot|$  est la norme euclidienne dans  $\mathbb{R}^3$  et  $\nabla \xi := (\nabla \xi_1, \nabla \xi_2, \nabla \xi_3)$ , pour tout  $\xi = (\xi_1, \xi_2, \xi_3) \in W_{\|\cdot\|}$ .

On peut vérifier aisément que la fonctionnelle  $J$  est de classe  $\mathcal{C}^1(W_{\|\cdot\|}, \mathbb{R})$  et 2-elliptique, c'est-à-dire qu'il existe une constante  $M > 0$  telle que

$$(J'(\sigma) - J'(\xi)) \cdot (\sigma - \xi) \geq M \|\sigma - \xi\|^2, \quad \forall \sigma, \xi \in W_{\|\cdot\|}.$$

Ainsi, la fonctionnelle  $J$  est strictement convexe et coercive sur  $W_{\|\cdot\|}$ . Comme l'espace  $W_{\|\cdot\|}$  est réflexif, les méthodes directes du calcul variationnel entraînent

que le problème (3.4) admet une solution unique.

D'autre part, étant donné que  $H^1(\Omega)$  est un espace de Hilbert tensoriel, puisque

$$H^1(\Omega) = \overbrace{a \bigotimes_{k=1}^3 H^1(\Omega_k)}^{\|\cdot\|},$$

alors l'espace  $W_{\|\cdot\|}$  l'est également, comme produit cartésien (fini) d'espaces de

Hilbert tensoriels. L'espace tensoriel algébrique associé à  $W_{\|\cdot\|}$  est  $W := \left[ a \bigotimes_{k=1}^3 H^1(\Omega_k) \right]^3$ .

Pour décrire la méthode PGD dans le cadre vectoriel, on introduit l'ensemble

$$\mathcal{R}_1(W) := \left[ \mathcal{R}_1 \left( a \bigotimes_{k=1}^3 H^1(\Omega_k) \right) \right]^3$$

des tenseurs élémentaires dans  $W$ .

On sait que  $\mathcal{R}_1 \left( a \bigotimes_{k=1}^3 H^1(\Omega_k) \right)$  est faiblement fermé dans  $H^1(\Omega)$ , par conséquent

$\mathcal{R}_1(W)$  est faiblement fermé dans  $W_{\|\cdot\|}$ . Donc les hypothèses  $(H_1) - (H_3)$  sont satisfaites dans le cadre du modèle du flot optique que nous considérons. Ainsi pour tout  $\sigma \in W_{\|\cdot\|}$ , le problème

$$\inf_{z \in \mathcal{R}_1(W)} J(\sigma + z) \quad (3.5)$$

admet au moins une solution.

**Lemme 3.2.1.** *Supposons qu'il existe  $\sigma \in W_{\|\cdot\|}$  vérifiant*

$$J(\sigma) \leq J(\sigma + z), \quad \forall z \in \mathcal{R}_1(W).$$

*Alors  $\sigma$  est l'unique solution de (3.4).*

**Preuve.** Soit  $z = (\varphi, \psi, \xi) \in \mathcal{R}_1(W)$  et soit la fonction  $g : \mathbb{R}^3 \rightarrow \mathbb{R}$  définie par  $g(r, s, t) = J(\sigma_1 + r\varphi, \sigma_2 + s\psi, \sigma_3 + t\xi)$ . Il est clair que la fonction  $g$  est différentiable sur  $\mathbb{R}^3$  et réalise son minimum absolue en  $(0, 0, 0)$ . Par conséquent  $g'(0, 0, 0) = J'(\sigma) \cdot (\varphi, \psi, \xi) = 0$ . Par linéarité de  $J'(\sigma)$ , on déduit que  $J'(\sigma) = 0$  sur  $W$ . La densité de  $W$  dans  $W_{\|\cdot\|}$  permet de conclure.  $\square$

Dans la suite, nous considérons la méthode PGD qui consiste en la stratégie (c) pour l'approximation progressive de la solution :

$$(M) \quad \begin{cases} (i) & \text{Initialisation : } (u_0, v_0, w_0) := (0, 0, 0) \text{ dans } W_{\|\cdot\|}. \\ (ii) & \text{Choisir } (\varphi_m, \psi_m, \xi_m) := \arg \min_{(\varphi, \psi, \xi) \in \mathcal{R}_1(W)} J(u_{m-1} + \varphi, v_{m-1} + \psi, w_{m-1} + \xi). \\ (iii) & (u_m, v_m, w_m) := (u_{m-1}, v_{m-1}, w_{m-1}) + (\varphi_m, \psi_m, \xi_m), \quad \forall m \geq 1. \end{cases}$$

Il s'agit de la méthode la plus utilisée en PGD, grâce à sa simplicité d'implémentation. Nous montrons que toute suite PGD de type (M) est convergente dans  $W_{\|\cdot\|}$ .

**Théorème 3.2.2.** Soit  $\Omega = \prod_{k=1}^3 \Omega_k$  un domaine de  $\mathbb{R}^3$ , où  $(\Omega_k)_{1 \leq k \leq 3}$  désignent des intervalles de  $\mathbb{R}$ , pas forcément bornés. Soit  $J$  la fonctionnelle définie sur  $W_{\|\cdot\|}$  par (3.3), où la fonction  $I \in W^{1,\infty}(]0, +\infty[ \times \Omega)$  et la fonction  $c$  est continue sur  $\Omega$ , positive et bornée. Alors la suite  $(u_m, v_m, w_m)_{m \in \mathbb{N}} \subset W$ , définie par (M), converge faiblement dans  $W_{\|\cdot\|}$  vers l'unique solution du problème (3.4).

**Preuve.** Soit une suite  $(u_m, v_m, w_m)_{m \in \mathbb{N}} \subset W$  de type PGD, définie par (M). Il est clair que la suite réelle  $(J(u_m, v_m, w_m))_{m \in \mathbb{N}}$  est décroissante, minorée et par conséquent convergente dans  $\mathbb{R}$ , on notera  $\ell$  sa limite. La coercivité de  $J$  entraîne que la suite  $(u_m, v_m, w_m)_{m \in \mathbb{N}}$  est également bornée dans  $W_{\|\cdot\|}$ .

Montrons d'abord que la suite  $(u_m, v_m, w_m)_m$  converge faiblement dans  $W_{\|\cdot\|}$ . Pour cela, montrons que la suite  $(u_m, v_m, w_m)_m$  a une unique valeur d'adhérence pour la topologie faible de  $W_{\|\cdot\|}$ .

Soit  $(u, v, w) \in W_{\|\cdot\|}$  une valeur d'adhérence de  $(u_m, v_m, w_m)_m$  pour la topologie faible de  $W_{\|\cdot\|}$  et soit  $(u_{m_k}, v_{m_k}, w_{m_k})_k$  une sous-suite de  $(u_m, v_m, w_m)_m$  convergeant faiblement vers  $(u, v, w)$ .

Comme  $u_{m_k} \rightharpoonup u$  dans  $H^1(\Omega_1)$  alors on a  $u_{m_k} + \varphi \rightharpoonup u + \varphi$ , pour tout  $\varphi \in H^1(\Omega_1)$ . Donc

$$\begin{aligned} \|u_{m_k} - u\|_{L^2(\Omega_1)}^2 &= \|u_{m_k} + \varphi\|_{L^2(\Omega_1)}^2 + \|u + \varphi\|_{L^2(\Omega_1)}^2 - 2 \langle u_{m_k} + \varphi, u + \varphi \rangle_{L^2(\Omega_1)} \\ &= \|u_{m_k} + \varphi\|_{L^2(\Omega_1)}^2 - \|u + \varphi\|_{L^2(\Omega_1)}^2 + o_k(1) \end{aligned} \quad (3.6)$$

et

$$\begin{aligned} \|\nabla u_{m_k} - \nabla u\|_{L^2(\Omega_1)}^2 &= \|\nabla u_{m_k} + \nabla \varphi\|_{L^2(\Omega_1)}^2 + \|\nabla u + \nabla \varphi\|_{L^2(\Omega_1)}^2 - 2 \langle \nabla(u_{m_k} + \varphi), \nabla(u + \varphi) \rangle_{L^2(\Omega_1)} \\ &= \|\nabla u_{m_k} + \nabla \varphi\|_{L^2(\Omega_1)}^2 - \|\nabla u + \nabla \varphi\|_{L^2(\Omega_1)}^2 + o_k(1). \end{aligned} \quad (3.7)$$

Ainsi, en utilisant les mêmes arguments pour les sous-suites  $(v_{m_k})_k$  et  $(w_{m_k})_k$ , on déduit que

$$J(u_{m_k} - u, v_{m_k} - v, w_{m_k} - w) = J(0) + J(u_{m_k}, v_{m_k}, w_{m_k}) - J(u, v, w) + o_k(1).$$

D'autre part, on a

$$J(u_{m_{k+1}}, v_{m_{k+1}}, w_{m_{k+1}}) \leq J(u_{m_k+1}, v_{m_k+1}, w_{m_k+1}) \leq J(u_{m_k} + \varphi, v_{m_k} + \psi, w_{m_k} + \xi), \quad \forall (\varphi, \psi, \xi) \in \mathcal{R}_1(W).$$

Ainsi, pour tout  $(\varphi, \psi, \xi) \in \mathcal{R}_1(W)$  on a

$$J(u_{m_{k+1}} - u, v_{m_{k+1}} - v, w_{m_{k+1}} - w) + J(u, v, w) \leq J(u_{m_k} - u, v_{m_k} - v, w_{m_k} - w) + J(u + \varphi, v + \psi, w + \xi) + o_k(1).$$

En remarquant que  $J(u_{m_{k+1}} - u, v_{m_{k+1}} - v, w_{m_{k+1}} - w)$  et  $J(u_{m_k} - u, v_{m_k} - v, w_{m_k} - w)$  ont la même limite  $\ell - J(u, v, w)$  lorsque  $k \rightarrow +\infty$ , on obtient par passage à la limite que

$$J(u, v, w) \leq J(u + \varphi, v + \psi, w + \xi), \quad \forall (\varphi, \psi, \xi) \in \mathcal{R}_1(W).$$

Le Lemme 3.2.1 entraîne que  $(u, v, w)$  est l'unique solution du problème (3.4). Ainsi, la suite  $(u_m, v_m, w_m)_{m \in \mathbb{N}}$  possède une unique valeur d'adhérence pour la topologie faible de  $W_{\|\cdot\|}$ , donc elle converge faiblement vers la solution du problème (3.4).  $\square$

**Corollaire 3.2.3.** *Sous les mêmes conditions que le Théorème 3.2.2, on a :*

$$J'(\sigma_m) \rightharpoonup 0 \quad \text{dans } W'_{\|\cdot\|} \text{ faible } \star.$$

*Autrement dit*

$$\lim_{m \rightarrow +\infty} J'(\sigma_m) \cdot z = 0, \quad \forall z \in \mathcal{R}_1(W).$$

**Preuve.** Il suffit d'écrire l'expression de  $J'(\sigma_m) \cdot z$  et utiliser la convergence faible de  $(\sigma_m)_m$  vers l'unique solution du problème (3.4).  $\square$

**Lemme 3.2.4.** *La suite des directions successives de descentes  $z_m := (\varphi_m, \psi_m, \xi_m)$  définie dans (M) pour la suite PGD  $(\sigma_m)_m$  vérifie :*

$$\sum_{k=1}^{+\infty} \|z_k\|^2 < +\infty,$$

*en particulier,  $\lim_{m \rightarrow +\infty} z_m = 0$  dans  $W_{\|\cdot\|}$ .*

**Preuve.** Introduisons la partie quadratique de  $J$  définie sur  $W_{\|\cdot\|}$  par :

$$\tilde{J}(\sigma) := J(\sigma) - J(0) - J'(0) \cdot \sigma.$$

Ainsi,

$$J'(\sigma) \cdot \chi = \tilde{J}'(\sigma) \cdot \chi + J'(0) \cdot \chi,$$

où la forme définie sur  $W_{\|\cdot\|} \times W_{\|\cdot\|}$  par :  $\chi \mapsto \tilde{J}'(\sigma) \cdot \chi$  est bilinéaire symétrique.

Il suit alors que

$$\begin{aligned}
 J'(\sigma_m) \cdot \sigma_m &= J'(\sigma_m) \cdot \sigma_{m-1} \\
 &= \tilde{J}'(\sigma_m) \cdot \sigma_{m-1} + J'(0) \cdot \sigma_{m-1} \\
 &= \tilde{J}'(\sigma_{m-1}) \cdot \sigma_{m-1} + J'(0) \cdot \sigma_{m-1} + \tilde{J}'(z_m) \cdot \sigma_{m-1} \\
 &= J'(\sigma_{m-1}) \cdot \sigma_{m-1} + \tilde{J}'(\sigma_{m-1}) \cdot z_m \\
 &\vdots \\
 &= \sum_{1 \leq i < j \leq m} \tilde{J}'(z_i) \cdot z_j \\
 &= \frac{1}{2} \left( \tilde{J}'(\sigma_m) \cdot \sigma_m - \sum_{k=1}^m \tilde{J}'(z_k) \cdot z_k \right) \\
 &= \frac{1}{2} \left( J'(\sigma_m) \cdot \sigma_m - J'(0) \cdot \sigma_m - \sum_{k=1}^m \tilde{J}'(z_k) \cdot z_k \right).
 \end{aligned}$$

D'où

$$J'(\sigma_m) \cdot \sigma_m = -J'(0) \cdot \sigma_m - \sum_{k=1}^m \tilde{J}'(z_k) \cdot z_k. \quad (3.8)$$

Comme  $\tilde{J}'(z_k) \cdot z_k \geq C_1 \|z_k\|^2$ , pour tout entier  $k$ , il s'ensuit que

$$\sum_{k=1}^m \|z_k\|^2 \leq C_2 (\|J'(\sigma_m)\|_* \times \|\sigma_m\| + \|J'(0)\|_* \times \|\sigma_m\|) \leq C_3.$$

En particulier,  $\lim_{m \rightarrow +\infty} \|z_m\| = 0$ . □

**Lemme 3.2.5.** *Soit  $(\sigma_m)_{m \in \mathbb{N}}$  une suite PGD pour le problème (3.4), alors il existe une constante  $C > 0$  telle que :*

$$|J'(\sigma_{m-1}) \cdot z| \leq C \|z_m\| \|z\|,$$

pour tout  $z \in \mathcal{R}_1(W)$ , où  $z_m := (\varphi_m, \psi_m, \xi_m)$  est la direction de descente à l'itération  $m \geq 1$ , définie dans (M).

**Preuve.** Commençons par noter que la fonctionnelle  $J' : W_{\|\cdot\|} \rightarrow W_{\|\cdot\|}^*$  est lipshitzienne. D'autre part, la convexité de  $J$  entraîne que :

$$J'(\sigma_{m-1} + z) \cdot (z_m - z) \leq J(\sigma_{m-1} + z_m) - J(\sigma_{m-1} + z) \leq 0,$$

ainsi

$$J'(\sigma_{m-1} + z) \cdot z_m \leq J'(\sigma_{m-1} + z) \cdot z. \quad (3.9)$$

$$\begin{aligned}
 -J'(\sigma_{m-1}) \cdot z &= (J'(\sigma_{m-1} + z) - J'(\sigma_{m-1})) \cdot z - J'(\sigma_{m-1} + z) \cdot z \\
 &\leq C\|z\|^2 - J'(\sigma_{m-1} + z) \cdot z \quad (J \text{ lipschitzienne}) \\
 &\leq C\|z\|^2 - J'(\sigma_{m-1} + z) \cdot z_m \quad (\text{d'après 3.9}) \\
 &= C\|z\|^2 - (J'(\sigma_{m-1} + z) - J'(\sigma_{m-1} + z_m)) \cdot z_m \\
 &\leq C(\|z\|^2 + \|z - z_m\| \|z_m\|) \\
 &\leq C(\|z\|^2 + \|z\| \|z_m\| + \|z_m\|^2).
 \end{aligned}$$

D'où, en remplaçant  $z$  par  $\pm z \frac{\|z_m\|}{\|z\|}$  dans ce qui précède, on obtient :

$$|J'(\sigma_{m-1}) \cdot z| \leq 3C\|z_m\| \|z\|. \quad \square$$

**Lemme 3.2.6.** Avec les mêmes notations que le Théorème 3.2.2, on a :

$$\lim_{m \rightarrow +\infty} J'(\sigma_m) \cdot \sigma_m = 0$$

et la suite PGD  $(\sigma_m)_m$  donnée par l'algorithme (M) converge vers la solution du problème (3.4).

**Preuve.** D'après (3.8), on a

$$J'(\sigma_m) \cdot \sigma_m = -J'(0) \cdot \sigma_m - \sum_{k=1}^m \tilde{J}'(z_k) \cdot z_k.$$

Comme  $\sum_{k=1}^{+\infty} \tilde{J}'(z_k) \cdot z_k < +\infty$  et  $\lim_{m \rightarrow +\infty} J'(0) \cdot \sigma_m = J'(0) \cdot \hat{\sigma}$ , où  $\hat{\sigma}$  est la limite faible de  $(\sigma_m)_{m \in \mathbb{N}}$  (et au même temps l'unique solution de (3.4) d'après le Théorème 3.2.2), alors la suite réelle  $(J'(\sigma_m) \cdot \sigma_m)_m$  est convergente. D'autre part, on a

$$\begin{aligned}
 |J'(\sigma_m) \cdot \sigma_m| &\leq \sum_{k=1}^m |J'(\sigma_m) \cdot z_k| \\
 &\leq C \sum_{k=1}^m \|z_{m+1}\| \|z_k\| \quad (\text{Lemme 3.2.5}) \\
 &\leq C (m \|z_{m+1}\|^2)^{1/2} \left( \sum_{k=1}^m \|z_k\|^2 \right)^{1/2} \quad (\text{Inégalité de Hölder}) \\
 &\leq C ((m+1) \|z_{m+1}\|^2)^{1/2} \left( \sum_{k=1}^{+\infty} \|z_k\|^2 \right)^{1/2}
 \end{aligned}$$

Finalement, étant donné que  $\sum_{k=1}^{+\infty} \|z_k\|^2 < +\infty$  alors il existe une sous-suite  $(z_{m_k})_k$  de  $(z_m)_m$  telle que  $\lim_{k \rightarrow +\infty} m_k \|z_{m_k}\|^2 = 0$ . D'où

$$\lim_{m \rightarrow +\infty} J'(\sigma_m) \cdot \sigma_m = \lim_{k \rightarrow +\infty} J'(\sigma_{m_k}) \cdot \sigma_{m_k} = 0.$$

Finalement, en utilisant

$$\begin{cases} \sigma_m & \rightharpoonup \sigma & \text{dans } W_{\|\cdot\|} \\ J'(\sigma_m) \cdot \sigma_m & \rightarrow 0 & \text{quand } m \rightarrow +\infty \end{cases}$$

et les formules de décomposition (3.6) et (3.7), on déduit que

$$J'(\sigma_m - \sigma) \cdot (\sigma_m - \sigma) \rightarrow 0 \quad \text{quand } m \rightarrow +\infty.$$

Par conséquent

$$J'(0) \cdot (\sigma_m - \sigma) + \tilde{J}'(\sigma_m - \sigma) \cdot (\sigma_m - \sigma) \rightarrow 0 \quad \text{quand } m \rightarrow +\infty.$$

D'où le résultat.  $\square$

### 3.2.2 Implementation des approches PGD et Horn&Schunck en flot optique

La méthode PGD est comparée à la technique de Horn&Schunck (HS) aussi bien au niveau du temps de calcul qu'au niveaux de la précision et l'exactitude du mouvement estimé. On utilise un schéma multi-résolution pour passer itérativement d'une image sous-échantillonnée  $32 \times 32$  jusqu'à résolution originale. La procédure d'implémentation est décrite, pour les deux approches (PGD) & (HS) dans le diagramme 3.1. La description de l'étape spécifique à chaque méthode (en gris) dans le diagramme 3.1 est détaillée dans la section 3.2.2 et 3.2.2 pour les méthodes PGD et HS respectivement.

#### La méthode PGD proposée

Notons tout d'abord que le terme

$$\int_{\Omega} c(x, y, z) |\sigma(x, y, z)|^2 dx dy dz$$

qui apparait dans l'énergie  $J$  permet d'assurer que  $J$  est bien coercive sur son espace fonctionnel. D'un point de vue de la simulation numérique, mises à part les situations où les images sont fortement bruitées, ce terme n'a pas un impact significatif sur les résultats numériques, par conséquent, il sera omis dans toute la partie étude numérique qui suivra.

Le système d'Euler-Lagrange associé au problème (3.4) est donné par :

$$\begin{cases} -\alpha \Delta u + I_x^2 u + I_x I_y v & = -I_x I_t \\ -\alpha \Delta v + I_x I_y u + I_y^2 v & = -I_y I_t \end{cases} \quad (3.10)$$

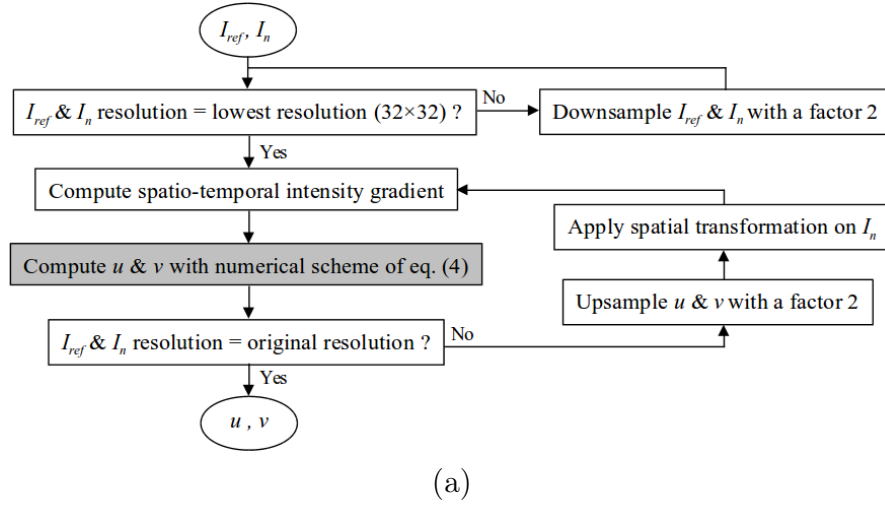


FIGURE 3.1 – Séquence de traitement de données pour la mise en oeuvre des approches  $L^2 - L^2$  du flot optique.

On cherche la solution de ce système de la manière itérative suivante :

$$\begin{cases} u_m &= u_{m-1} + \varphi_m \otimes \psi_m \\ v_m &= v_{m-1} + \widetilde{\varphi}_m \otimes \widetilde{\psi}_m \end{cases} \quad (3.11)$$

avec

$$\begin{cases} u_{m-1} &= \sum_{k=1}^{m-1} \varphi_k \otimes \psi_k \\ v_{m-1} &= \sum_{k=1}^{m-1} \widetilde{\varphi}_k \otimes \widetilde{\psi}_k \end{cases} \quad (3.12)$$

Pour chaque itération, les composantes de la PGD  $(\varphi_m, \widetilde{\varphi}_m, \psi_m, \widetilde{\psi}_m)$  sont solutions du système non-linéaire formé de  $\text{Card}(\Omega_1)^2 \times \text{Card}(\Omega_2)^2$  équations :

$$\begin{cases} -\alpha \varphi_m'' \psi_m - \alpha \varphi_m \psi_m'' + I_x^2 \varphi_m \psi_m + I_x I_y \widetilde{\varphi}_m \widetilde{\psi}_m &= -I_x I_t + \alpha \Delta u_{m-1} \\ -\alpha \widetilde{\varphi}_m'' \widetilde{\psi}_m - \alpha \widetilde{\varphi}_m \widetilde{\psi}_m'' + I_x I_y \varphi_m \psi_m + I_y^2 \widetilde{\varphi}_m \widetilde{\psi}_m &= -I_y I_t + \alpha \Delta v_{m-1} \end{cases} \quad (3.13)$$

Pour résoudre (3.13), on procède en deux étapes, une première pour calculer  $(\varphi_m, \widetilde{\varphi}_m)$  et une seconde pour calculer  $(\psi_m, \widetilde{\psi}_m)$ .

Pour calculer  $(\varphi_m, \widetilde{\varphi}_m)$ , on multiplie les deux premières équations de (3.13) par  $\psi_m$  et  $\widetilde{\psi}_m$  respectivement ensuite on intègre le résultat obtenu par rapport



à la variable  $y$  pour aboutir au système de  $2 \times \text{Card}(\Omega_1)$  équations d'inconnues  $\varphi_m(x)$  et  $\widetilde{\varphi}_m(x)$ .

$$\begin{cases} A\varphi_m''(x) + B(x)\varphi_m(x) + C(x)\widetilde{\varphi}_m(x) = D(x) \\ \widetilde{A}\varphi_m''(x) + \widetilde{B}(x)\varphi_m(x) + \widetilde{C}(x)\widetilde{\varphi}_m(x) = \widetilde{D}(x) \end{cases} \quad (3.14)$$

où

$$\begin{cases} A = -\alpha \int_{\Omega_2} \psi_m(y)^2 dy \\ B(x) = \int_{\Omega_2} \psi_m'^2(y) dy + \int_{\Omega_2} I_x(x, y)^2 \psi_m^2(y) dy \\ C(x) = \int_{\Omega_2} I_x(x, y) I_y(x, y) \widetilde{\psi}_m(y) \psi_m(y) dy \\ D(x) = - \int_{\Omega_2} I_y(x, y) I_y(x, y) \psi_m(y) dy + \alpha \int_{\Omega_2} \Delta u_{m-1}(x, y) \psi_m(y) dy \end{cases} \quad (3.15)$$

On utilise l'approximation du Laplacien suggérée par Horn&Schunk, où  $\Delta u \simeq \overline{u_k} - u_k$ ,  $\overline{u_k}$  étant la moyenne de  $u_k$  sur un voisinage de  $3 \times 3$  pixels du point considéré []. La même approximation est utilisée pour le calcul de  $\Delta v$ .

De manière analogue, pour l'approximation unidimensionnelle de  $\varphi''$ , on utilise  $\varphi'' \simeq \overline{\varphi_k} - \varphi_k$ ,  $\overline{\varphi_k}$  étant la moyenne de  $\varphi_k$  sur un voisinage de 3 pixels du point considéré. La même approximation est utilisée pour le calcul de  $\varphi''$ ,  $\psi''$  et  $\psi''$ .

La solution  $(\varphi_m, \widetilde{\varphi}_m)$  du système (3.14) peut être obtenu par le schéma de point fixe :

$$\begin{cases} \varphi_m^{k+1} = \frac{(D - A\overline{\varphi^k})(\widetilde{C} - \widetilde{A}) - (\widetilde{D} - \widetilde{A}\overline{\varphi^k})C}{(B - A)(\widetilde{C} - \widetilde{A}) - \widetilde{B}C} \\ \widetilde{\varphi}_m^{k+1} = \frac{(B - A)(\widetilde{D} - \widetilde{A}\overline{\varphi^k}) - \widetilde{B}(D - A\overline{\varphi^k})}{(B - A)(\widetilde{C} - \widetilde{A}) - \widetilde{B}C} \end{cases} \quad (3.16)$$

La variation de  $\|(\varphi_m^k, \widetilde{\varphi}_m^k)\|_2$  entre deux itérations successives est comparée à un coefficient de tolérance maximal de  $10^{-6}$  pour assurer la convergence.

A ce stade, seulement  $2 \times \text{Card}(\Omega_1)$  équations ont été résolues pour obtenir  $(\varphi_m, \widetilde{\varphi}_m)$ .

Ensuite  $(\psi_m, \widetilde{\psi}_m)$  est obtenu de manière similaire en utilisant cependant la valeur de  $(\varphi_m, \widetilde{\varphi}_m)$  mise à jour, de la manière suivante :

$$\begin{cases} \psi_m^{k+1} = \frac{(H-E\overline{\psi^k})(\tilde{G}-\tilde{E}) - (\overline{H}-\tilde{E}\overline{\psi^k})G}{(F-E)(\tilde{G}-\tilde{E}) - \tilde{F}G} \\ \widetilde{\psi_m^{k+1}} = \frac{(F-E)(\tilde{H}-\tilde{E}\overline{\psi^k}) - \tilde{F}(H-E\overline{\psi^k})}{(F-E)(\tilde{G}-\tilde{E}) - \tilde{F}G} \end{cases} \quad (3.17)$$

où

$$\begin{cases} E &= -\alpha \int_{\Omega_1} \varphi_m(x)^2 dx \\ F(x) &= \int_{\Omega_1} \varphi_m'^2(x) dx + \int_{\Omega_2} I_x(x, y)^2 \varphi_m^2(x) dx \\ G(x) &= \int_{\Omega_1} I_x(x, y) I_y(x, y) \widetilde{\varphi_m}(y) \varphi_m(x) dx \\ H(x) &= - \int_{\Omega_1} I_y(x, y) I_y(x, y) \varphi_m(x) dx + \alpha \int_{\Omega_2} \Delta v_{m-1}(x, y) \varphi_m(x) dx \end{cases} \quad (3.18)$$

Ainsi, un second ensemble de  $2 \times \text{Card}(\Omega_2)$  équations a été résolu pour obtenir  $(\psi_m, \widetilde{\psi_m})$ .

Le processus global est stoppé une fois la variation  $\left\| \widetilde{\varphi_m} \otimes \widetilde{\psi_m} \right\|_2$  est inférieure à un coefficient de tolérance  $\varepsilon$  défini par l'utilisateur.

Pour la condition de convergence, l'erreur quadratique moyenne est calculée entre les champs de mouvements de l'itération courante et le champ de mouvement de l'itération précédente, puis elle est comparée à un epsilon (qui est le même pour Horn-Schunck et pour la PGD). Le choix du epsilon est implémenté de cette façon : pour résoudre le problème à la résolution originale, on résout notre énergie à une résolution basse (ici, on commence à 32x32 pixels), puis on résout à nouveau notre énergie à une résolution deux fois plus élevées et ainsi de suite, jusqu'à la résolution originale. Afin d'être très précis, on utilise un epsilon très fin à de basses résolutions, car leur impact est très important pour les hautes résolutions, et le epsilon est moins fin à haute résolution. Pour la résolution originale nous prenons un epsilon de 0.01 pixels. En pratique, la précision du mouvement réel par rapport au mouvement simulé est toujours très supérieure à 0.01, et cette valeur donne de bons résultats. Puis, pour chaque niveau de multi-résolution, le epsilon varie linéairement en fonction de la résolution de la manière suivante :

pour une image originale de 128x128, nous résolvons l'énergie pour les problèmes suivants :

- la résolution 32x32 avec un epsilon de 0.01 divisé par 4

- la résolution 64x64 avec un epsilon de 0.01 divisé par 2
- la résolution 128x128 avec un epsilon de 0.01

### L'approche de Horn&Schunck

Le système d'équations d'Euler-Lagrange (3.10) est résolu par un schéma itératif utilisant la méthode de Jacobi, comme suit :

$$\begin{cases} u^{k+1} = \overline{u^k} - \frac{I_x \cdot \overline{u^k} + I_y \cdot \overline{v^k}}{\alpha^2 + I_x^2 + I_y^2} \\ v^{k+1} = \overline{v^k} - \frac{I_x \cdot \overline{u^k} + I_y \cdot \overline{v^k}}{\alpha^2 + I_x^2 + I_y^2} \end{cases} \quad (3.19)$$

A ce stade,  $\text{Card}(\Omega_1) \times \text{Card}(\Omega_2)$  sont requises pour obtenir  $u$  et  $v$ .

La variation entre deux itérations successives est comparée à un coefficient de tolérance maximal de  $10^{-6}$  pour assurer la convergence. Pour mettre en évidence l'avantage de l'une des deux méthodes sur l'autre au niveau du résultat final, nous avons utilisé des schémas multi-résolution identiques pour les deux approches. Les conditions aux limites sont identiques, à savoir des conditions aux limites de type Neumann homogènes. Le calcul des dérivées spatio-temporelles et l'approximation du Laplacien sont également les mêmes. On réfère le lecteur à la section 3.2.2 pour plus de détails.

En outre, un raffinement itératif est utilisé pour mettre à jour  $I_{x,y,t}$  à l'intérieur de chaque résolution. Pour assurer la convergence de l'algorithme, la variation  $\left\| \overrightarrow{(u,v)} \right\|_2$  est comparée à un coefficient de tolérance maximal  $\varepsilon$  (voir la section 3.2.2 pour le choix de  $\varepsilon$ ).

### 3.2.3 Simulations numériques

La capacité de l'approche PGD pour calculer le flot optique, dans le cas d'énergie de type  $L^2 - L^2$ , a été évaluée en utilisant les scénarios suivants :

**Test de fiabilité #1 - Capacité de la méthode PGD à estimer une transformation tensorielle :** Une translation de 1 pixel suivant les directions horizontale et verticale a été appliquée à une image de référence pour obtenir l'image déplacée. L'image de "Lena" a été utilisée avec une résolution de  $512 \times 512$  pixels.

**Test de fiabilité #2 - Capacité de la méthode PGD à estimer une transformation non tensorielle :** Une rotation globale de 2 degré, centrée au milieu de l'image, a été appliquée à l'image de "Lena" pour obtenir l'image déplacée.

**Test de fiabilité #3 - Comportement de la méthode PGD vis à vis de la perturbation du niveau de gris :** Les images concernées dans le test de fiabilité #2 ont subi une perturbation par addition d'un bruit blanc gaussien d'écart-type  $\sigma = 10\%$  de l'écart maximal dans l'image du niveau de gris.

**Test de fiabilité #4 - Capacité de la méthode PGD à estimer une transformation locale de l'image :** Une image  $512 \times 512$  contenant 200 centroïdes (forme gaussienne  $\sigma = 2$  pixels) ont été générées. Chaque centroïde a subi un décalage aléatoire (le décalage horizontal et vertical suit une distribution gaussienne d'écart-type  $\sigma = 2$  pixels) pour obtenir l'image déplacée.

**Test de fiabilité #5 - Illustration pratique de l'avantage de la méthode PGD en imagerie médicale :** L'acquisition des deux images successives se fait par IRM (Imagerie par résonance magnétique) sur l'abdomen d'un volontaire sain à deux instants opposés du cycle respiratoire. Pour cela, deux images coronales ont été obtenues sur une machine IRM de type Tesla Philips Achieva 1.5 avec une séquence d'échos ( $TE = 18\text{ms}$ ). Une seule tranche, de résolution  $128 \times 128$  avec un voxel de  $2.5 \times 2.5 \times 8 \text{ mm}^3$ , est acquise en 300 ms. Ainsi, les deux reins, le foie et la rate sont observés dans le champs de vision des deux images.

### 3.2.4 Evaluation de la méthode PGD proposée

Pour les tests #1, #2 and #3, la distance euclidienne entre le déplacement estimé et celui simulé a été calculée. Pour le test #4, la distance euclidienne entre les décalages simulés et estimés est calculée aux centres des centroïdes. Comme le mouvement réel n'est pas disponible dans le cas du scénario #5, la distance euclidienne calculée est entre les champs de déplacement obtenus par les deux approches HS et PGD. Il est important que dans les tests #1, #2 and #3, la transformation spatiale simulée a été appliquée après l'étape de rééchantillonnage. Ainsi, pour le test #1, un décalage de  $\sqrt{2}$  pixels doit être estimé à chaque niveau de résolution. Au contraire, pour les tests #2 à #5, l'amplitude du mouvement simulé le long de des directions horizontale et verticale augmente linéairement en fonction de la dimension de l'image.

Pour étudier l'impact de la résolution de l'image sur la performance de la méthode PGD proposée, nous avons augmenté la résolution des images par interpolation bi-cubique. Ainsi, les résolutions suivantes ont été obtenues :  $64 \times 64$ ,  $128 \times 128$ ,  $256 \times 256$ ,  $512 \times 512$ ,  $1024 \times 1024$  et  $2048 \times 2048$ .

## 3.3 Résultats

La Figure 3.2 décrit des conclusions typiques de l'efficacité de l'approche PGD obtenues dans le cas d'un mouvement de translation simulé. La distribution pixel-

par-pixel de l'amplitude du mouvement est décrite pour chaque résolution (3.2b) : comme mentionné dans la section 3.2.4, un décalage commun de  $\sqrt{2}$  pixels est à estimer dans le cas de chaque résolution. La distribution pixel-par-pixel de l'erreur sur l'amplitude du mouvement pour la technique HS est représentée dans (3.2c) et pour l'approche PGD dans (3.2d), pour chaque résolution testée. On peut observer que l'erreur de la technique HS dépasse 0.1 pixel au-delà de la résolution  $512 \times 512$  pixels. Par contre, une erreur commune de l'ordre de 0.02 pixel est à noter, indépendamment de la résolution considérée. La carte d'erreur pixel-par-pixel pour une résolution de  $2048 \times 2048$  pixels est présentée dans la Figure 3.2e) et 3.2f) pour les approches HS et PGD respectivement. Des taux erreurs élevés ont lieu dans les régions homogènes de l'image étudiée (3.2e). La dernière carte d'erreurs est présentée dans 3.2a). La figure 3.2h) montre que le temps de calcul de la PGD est toujours inférieur à celui de la technique HS, en particulier pour les hautes résolutions ( $\approx 52$  s nécessaires pour une résolution de  $2048 \times 2048$  pixels pour la technique HS contre  $\approx 4$  s pour la PGD).

L'analyse réalisée, utilisant une transformation rotationnelle simulée et, les résultats, ont été reportés dans Fig. 3.3. Dans ce scénario, l'amplitude de la transformation simulée a augmenté, avec la dimension de l'image employée (voir 3.3b). Cette tendance avait un impact direct sur l'erreur de l'estimation du mouvement, qui à son tour, a augmenté de manière similaire, avec la résolution d'image (voir 3.3c et 3.3d). Des erreurs comparables sur les estimations de mouvement ont été obtenues par les approches HS (3.3c) et PGD (3.3d). Les cartes d'erreur de pixel par pixel obtenus pour une résolution de  $2048 \times 2048$  pixels (rapporté dans 3.3e et 3.3f) montre que la précision au centre de l'image a été améliorée en utilisant l'approche HS, par rapport à l'approche PGD. D'autre part, l'approche PGD a surpassé HS aux bords de l'image. La référence de la figure 3.3h montre que, une fois encore, le temps de calcul qui était obligatoire en utilisant l'approche PGD, a été réduit en comparaison avec l'approche HS, surtout pour les niveaux de haute résolution. Presque 160 s étaient obligatoires pour les images  $2048 \times 2048$  en utilisant l'approche HS, contre environ 10 s en utilisant la PGD.

La figure 3.4 montre le comportement de la méthode PGD sur des images soumises à des perturbations d'intensité de niveau de gris. De la même manière que le test de fiabilité #2, des erreurs d'estimation de mouvements comparables ont été obtenues pour les deux approches HS (3.4c) et PGD (3.4d), avec cependant, un léger avantage cette fois-ci, pour la PGD. Par rapport au test de fiabilité #2, la méthode HS avait un peu plus de difficulté à estimer les déplacements dans le centre de l'image (3.3f). L'indice de référence de la figure 3.3h montre que le temps de calcul avec la méthode PGD est très réduit par rapport au temps de calcul nécessaire avec la méthode HS, spécialement, aux niveaux des hautes résolutions (environ 600 s et 10 s ont été requis pour des images  $2048 \times 2048$  en utilisant respectivement les méthodes HS et PGD).

La figure 3.5 évalue la capacité de la méthode PGD pour estimer une transfor-

mation locale de l'image. De nouveau, la PGD a fournit des erreurs d'estimation de mouvements par rapport à l'approche HS (voir 3.5c et 3.5d) au niveau d'un temps de calcul réduit, plus faible, spécialement pour des images à haute résolution (voir 3.5f). Il est intéressant de noter que, jusqu'à 50 composants PGD étaient nécessaires pour caractériser la déformation simulée (voir 3.5e).

La figure 3.5 donne une illustration pratique de l'avantage de l'approche PGD sur une application d'imagerie médicale. Nous rappelons que deux images RM ont été prises à deux instants opposés du cycle respiratoire (voir l'image de référence dans 3.5a, et la différence entre l'image à enregistrer et l'image de référence dans 3.5b).

Dans l'image de différence, la direction de déplacement est essentiellement verticale, comme prévu, avec une valeur plus élevée pour le foie et la rate par rapport aux reins. Par rapport à l'amplitude de mouvement estimé (3.5c), la différence entre les estimations obtenues, en utilisant les méthodes HS et PGD, a été réduite pour toutes les résolutions testées (3.5d). Jusqu'à 30 composants PGD étaient obligatoires pour les niveaux les plus bas de résolution, impliqués dans l'approche multi-résolution(3.5e). Comme précédemment, l'approche PGD avait besoin de moins de temps que l'approche HS, spécialement pour les niveaux de haute résolution (environ 73 s et 12 s était requises pour des images  $2048 \times 2048$  en utilisant respectivement les approches HS et PGD).

### 3.4 Discussion

En utilisant la méthode HS, un grand nombre d'itérations est nécessaire pour propager le mouvement dans les régions homogènes, particulièrement pour les hautes résolutions d'image, comme indiqué dans la figure 3.2e. Ce n'est pas le cas de l'approche de PGD, qui est intrinsèquement basée sur une approche plus globale : Seulement 3-4 composantes PGD sont par conséquent nécessaire pour l'estimation de la transformation séparable simulée dans le test de fiabilité #1 (voir Figure 3.2g). La stratégie employée par la méthode HS dans le test de fiabilité #2 consiste à propager itérativement des estimations de mouvement du centre vers les frontières de l'image. La méthode HS ainsi fonctionne avec succès, particulièrement dans le centre d'image, requiert peu d'itérations pour avoir une estimation précise du mouvement.

En comparaison, la méthode PGD, qui utilise une approche plus globale, fournit une précision réduite du mouvement estimé au centre de l'image. Des observations opposées ont été cependant réalisées au voisinage de la frontière de l'image. Globalement, on obtient des résultats comparables en utilisant les deux méthodes pour le test de fiabilité #2. Il est intéressant de noter que la méthode PGD a requis un nombre élevé de composantes pour la caractérisation d'un mouvement rotationnel (l'image ref fig :Rotation g) contrairement à un mouvement de translation (Figure 3.2g). Ceci n'avait cependant aucun impact notable sur

la performance de la PGD en terme de temps de calcul, comme indiqué dans la Figure 3.3h : alors qu'un grand nombre de composantes PGD étaient nécessaires pour des basses résolutions (qui requièrent intrinsèquement moins de temps de calcul), le nombre de composantes PGD a été réduit pour les hautes résolutions (qui demandent plus de temps de calculs). Des observations identiques ont été aussi réalisées concernant l'estimation de déformations plus complexes (voir Figures 3.5 et 3.6).

Il doit être indiqué également que pour la technique HS, les perturbations dues au bruit sont attribués à des petits déplacements locaux. Ceci a considérablement augmenté le temps de calcul en présence de bruit dans l'image (près de 600 s sont nécessaires, comme indiqué dans l'image 3.4 h) contrairement au scénario sans bruit ( $\approx 160$  s, comme indiqué l'image 3.3 h). Quant à l'approche PGD, il peut être observé que les temps de calcul sont comparables en l'absence ou en présence de bruit (environ 10 s). En général, lors de tous les scénarios testés, l'exactitude des résultats de l'approche PGD sont comparables voire meilleurs que ceux de la technique HS avec un temps de calcul beaucoup plus bas, en particulier pour les images haute résolution.

Les méthodes variationnelles présentent un grand avantage pour la segmentation d'image, le filtrage, l'estimation de mouvement et également en compression d'acquisition. Nous pensons que la méthode PGD peut présenter un énorme intérêt dans les domaines où le traitement rapide de grosses données est nécessaire. En guise d'exemple pratique, nous avons illustré l'approche PGD dans le domaine de l'imagerie médicale, Fig.3.6. Des développements récents en IRM rapide, associés aux stratégies de traitement de données rapides, permettent maintenant d'acquérir des informations fonctionnelles et positionnelles en temps réel pendant une procédure interventionnelle.

En thérapie du cancer [1], une stratégie d'estimation de mouvement précise et rapide est par conséquent fondamentale. Par exemple, pour les cancers abdominaux, ceci permet d'assurer une dose de traitement optimale, associée à une toxicité limitée pour les organes voisins à risque, lors de l'utilisation d'ultrasons concentrés et de haute intensité [2] ou la thérapie de radio [3]. L'approche de PGD proposée est ainsi une excellente alternative pour l'estimation rapide de mouvement d'organes dans ce contexte.

## 3.5 Conclusion

\* — \* — \* — \* — \* — \* — \* — \* — \*

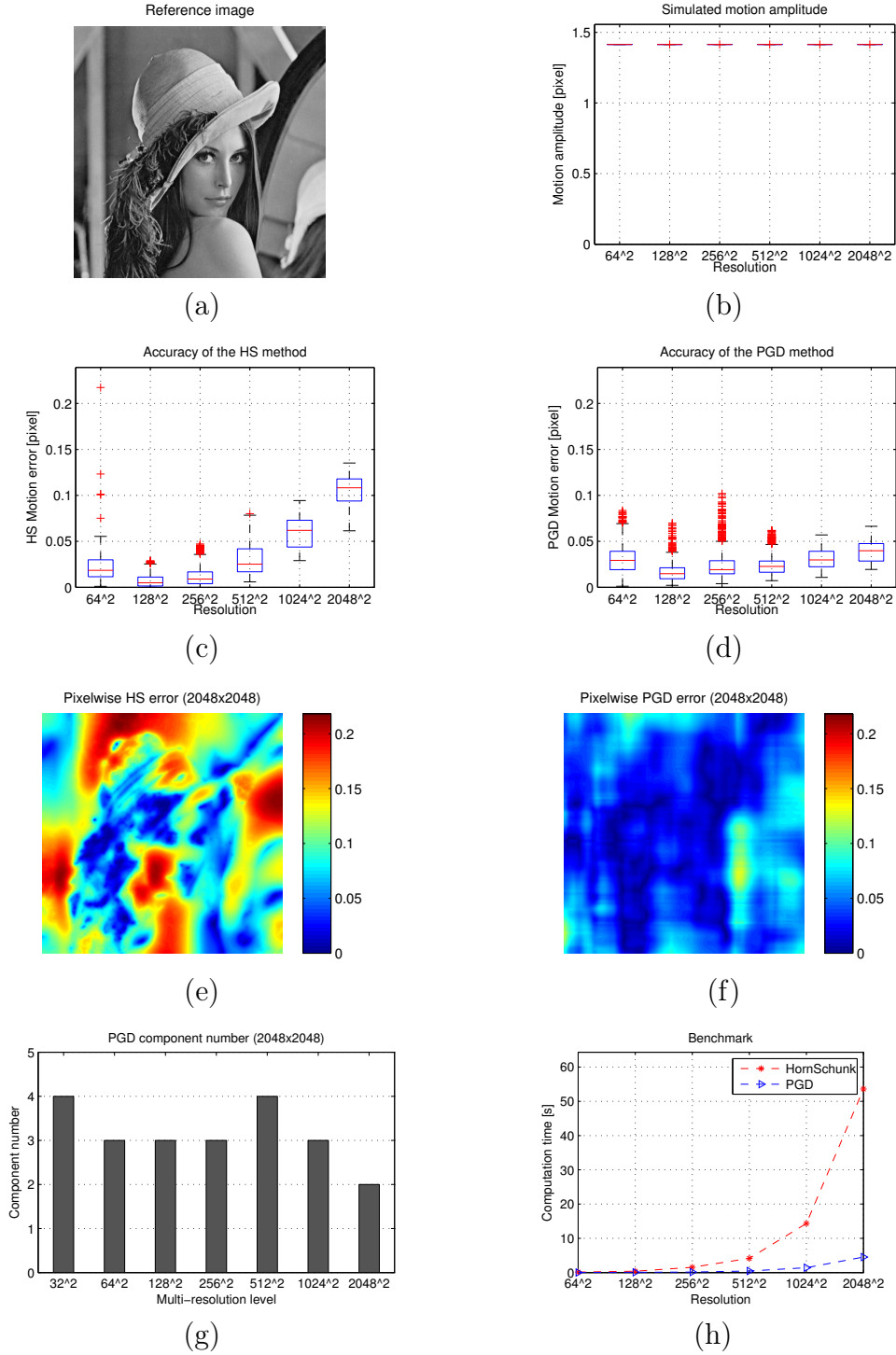


FIGURE 3.2 – Résultats typiques de l'efficacité de la PGD-basée Flot Optique obtenue sur une transformation de translation simulée (test de fiabilité #1). (a) : Image à enregistrer. (b) : Distribution de l'amplitude du champ de vecteurs de mouvement simulé pour chaque résolution testée. La médiane est représentée par la marque centrale, le premier et le troisième quartiles sont rapportés par les bords de la boîte, les moustaches s'étendent jusqu'aux valeurs les plus extrêmes qui ne sont pas considérées comme des valeurs aberrantes, et les valeurs aberrantes sont représentées individuellement en rouge. (c-d) : Distribution de l'erreur des estimations du mouvement obtenue pour chaque résolution testée en utilisant respectivement les approches HS (c) et PGD 63 proposées (d). (e-f) : L'erreur pixel à pixel des estimations du mouvement obtenue pour une résolution de  $2048 \times 2048$  pixels est rapportée pour les méthodes HS (e) et PGD (f). (g) : Nombre requis de composants PGD à chaque niveau de l'approche multi-résolution pendant l'estimation du flot optique d'une image  $2048 \times 2048$ . (h) : Le temps de calcul



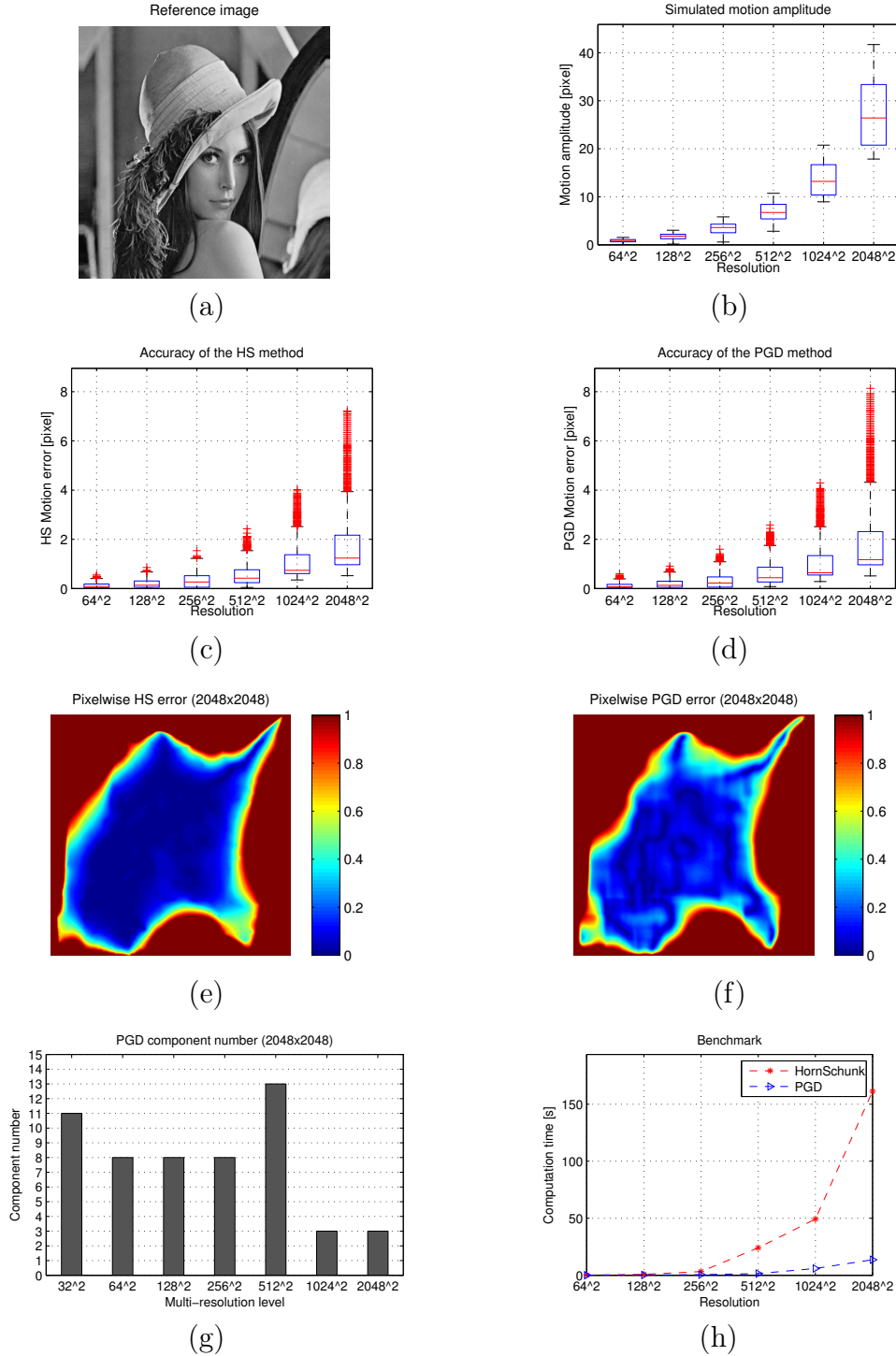


FIGURE 3.3 – Résultats typiques de l’efficacité de la PGD-basée Flot Optique obtenue sur une transformation de rotation simulée (test de fiabilité #2). (a) : Image à enregistrer. (b) : Distribution de l’amplitude du champ de vecteurs de mouvement simulé pour chaque résolution testée. (c-d) : Distribution de l’erreur des estimations du mouvement obtenue pour chaque résolution testée en utilisant respectivement les approches HS (c) et PGD proposées (d). (e-f) : L’erreur pixel à pixel des estimations du mouvement obtenue pour une résolution de  $2048 \times 2048$  pixels est rapportée pour les méthodes HS (e) et PGD (f). (g) : Nombre requis de composants PGD à chaque niveau de l’approche multi-résolution pendant l’estimation du flot optique d’une image  $2048 \times 2048$ . (h) : Le temps de calcul requis pour la mise en oeuvre des méthodes HS et PGD pour chaque résolution testée.

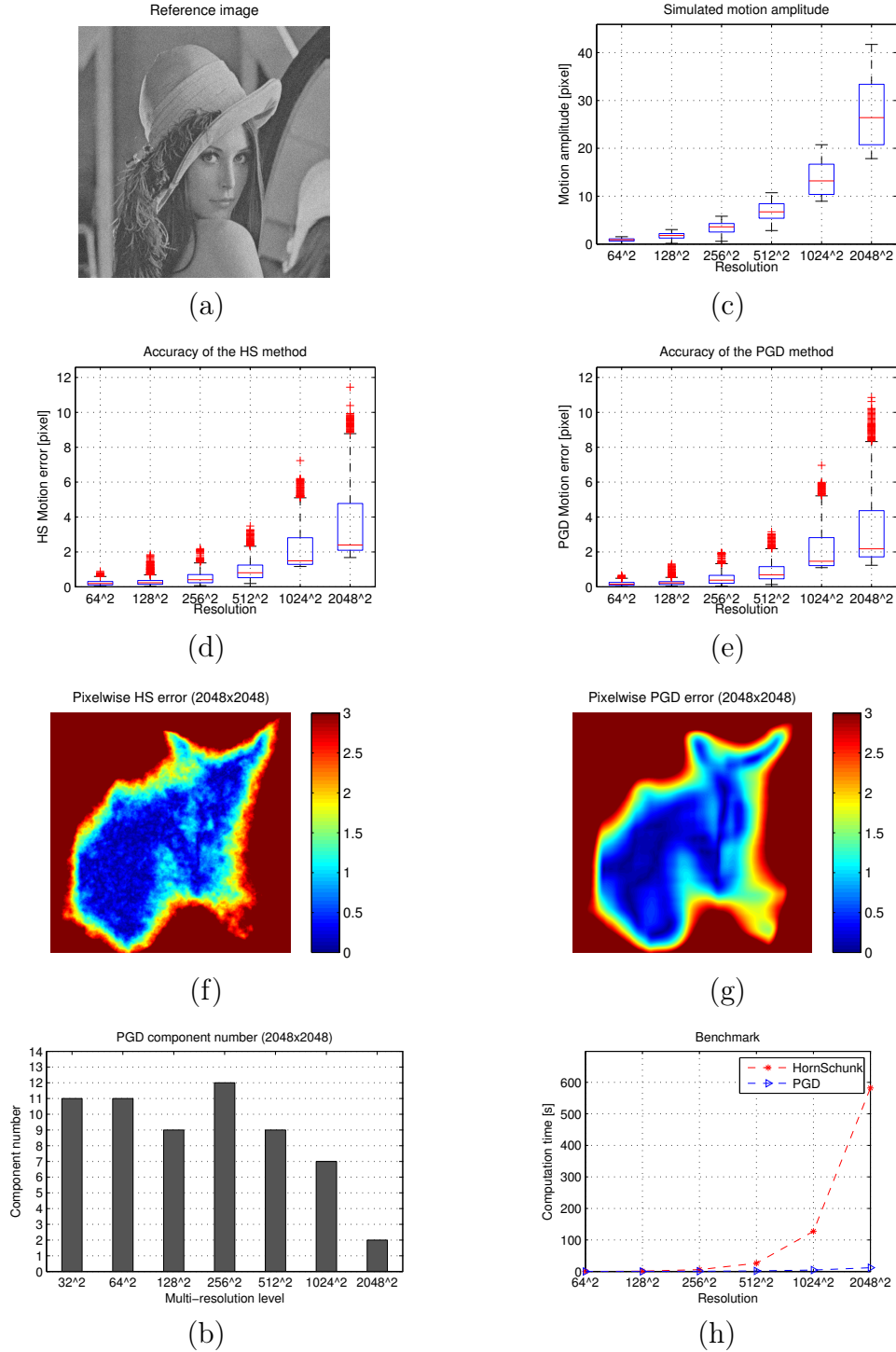


FIGURE 3.4 – Résultats typiques de l'efficacité de la PGD-basée Flot Optique obtenus sur une transformation de rotation simulée avec un bruit blanc gaussien supplémentaire (test de fiabilité #3). (a) L'image à enregistrer. (b) : Distribution de l'amplitude du champ de vecteurs de mouvement simulé pour chaque résolution testée. (c-d) : Distribution de l'erreur des estimations du mouvement obtenue pour chaque résolution testée en utilisant respectivement les approches HS (c) et PGD proposées (d). (e-f) : L'erreur pixel à pixel des estimations du mouvement obtenue pour une résolution de  $2048 \times 2048$  pixels est rapportée pour les méthodes HS (e) et PGD (f). (g) : Nombre requis de composants PGD à chaque niveau de l'approche multi-résolution pendant l'estimation du flot optique d'une image  $2048 \times 2048$ . (h) : Le temps de calcul requis pour la mise en oeuvre des méthodes HS et PGD pour chaque résolution testée.

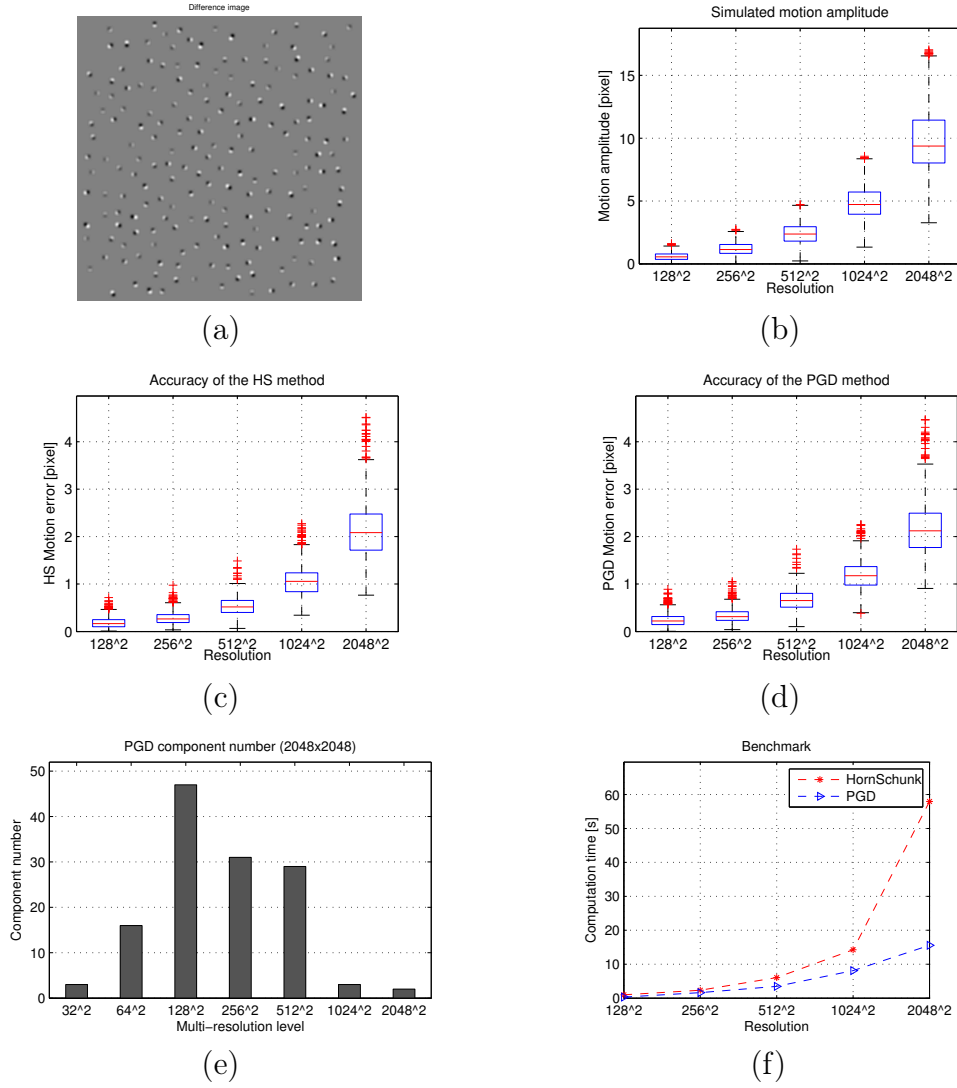


FIGURE 3.5 – Résultats typiques de l’efficacité de la PGD-basée Flot Optique obtenus dans un mouvement complexe de l’image locale (test de fiabilité #4). (a) : Différence entre l’image à enregistrer et l’image de référence. (b) : Distribution de l’amplitude du champ de vecteurs de mouvement simulé pour chaque résolution testée. (c-d) : Distribution de l’erreur des estimations du mouvement obtenue pour chaque résolution testée en utilisant respectivement les approches HS (c) et PGD proposées (d). (e) : Le nombre de composants PGD requis à chaque niveau de l’approche multi-résolution durant l’estimation du flot optique d’une image  $2048 \times 2048$ . (f) : Le temps de calcul requis pour la mise en oeuvre des méthodes HS et PGD pour chaque résolution testée.

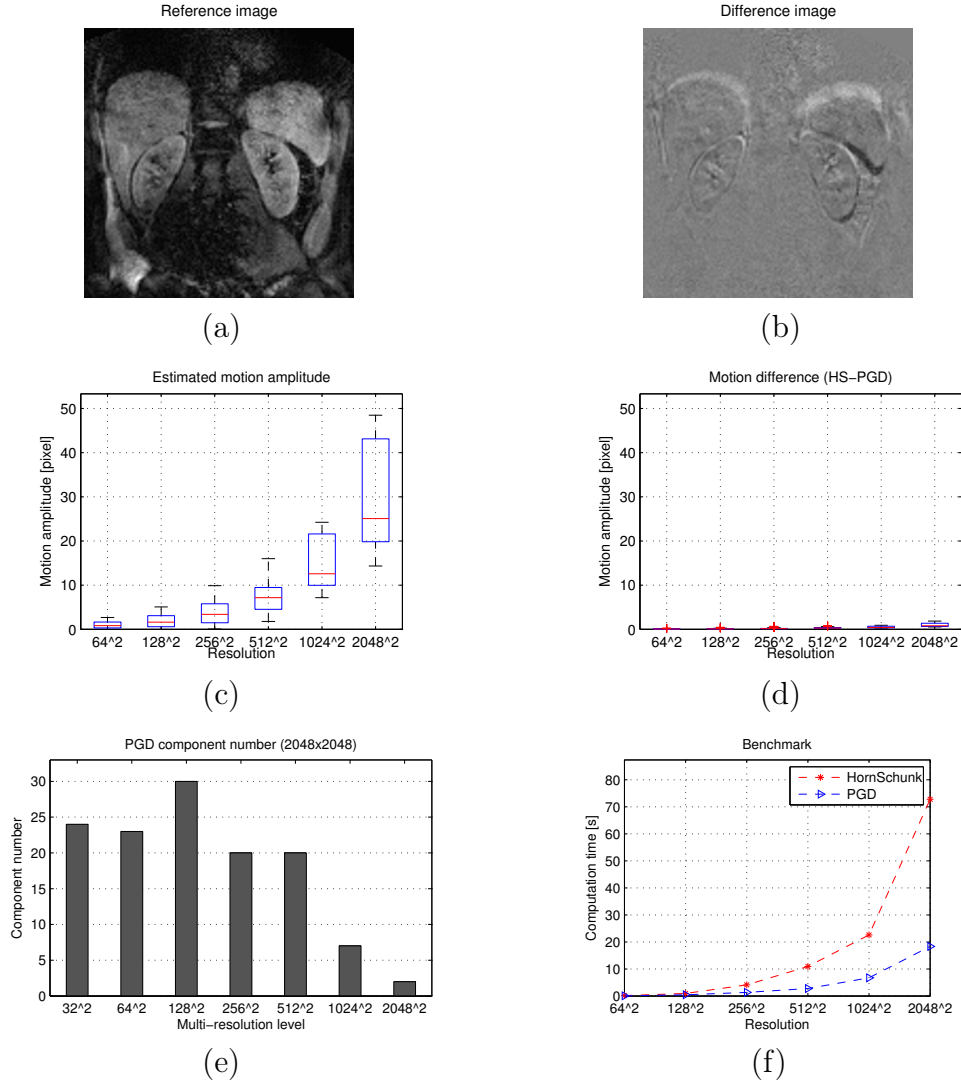


FIGURE 3.6 – Résultats typiques de l’efficacité de la PGD-basée Flot Optique obtenus pour des images RM abdominaux d’un volontaire sain (test de fiabilité #5). (a) : Image RM de référence. Les deux reins sont positionnés sous le foie (à gauche) et la rate (à droite). (b) : Différence entre l’image à enregistrer et l’image de référence. (c) : Distribution de l’amplitude du champ de mouvement simulé pour chaque résolution testée. (d) : Distribution de la différence entre les champs de mouvement obtenus en utilisant les méthodes HS et PGD. (e) : Le nombre de composants PGD requis à chaque niveau de l’approche multi-résolution durant l’estimation du flot optique d’une image  $2048 \times 2048$ . (f) : Le temps de calcul requis pour la mise en oeuvre des méthodes HS et PGD pour chaque résolution testée.



# Chapitre 4

## PGD sur la variété de Nehari pour les problèmes concave-convexe

### Sommaire

---

Présentation du chapitre . . . . .	70
4.1 Introduction . . . . .	70
4.2 Résultats préliminaires . . . . .	71
4.3 Résultats d'existence . . . . .	76
4.4 Decomposition Progressive Généralisée (PGD) . . . . .	77
4.5 Résultats numériques . . . . .	79

---

## 4.1 Introduction

Dans ce chapitre, nous traitons l'existence de solutions multiples pour le problème aux limites :

$$\begin{cases} -\Delta_p u = \lambda |u|^{q-2} u + |u|^{r-2} u & \text{dans } \Omega, \\ u = 0, & \text{sur } \partial\Omega, \end{cases} \quad (4.1)$$

par rapport au paramètre réel  $\lambda$ . Ici,  $\Omega$  est un domaine borné dans  $\mathbb{R}^N$ ,  $N \geq 2$ ,  $\partial\Omega$  son bord,  $\Delta_p$  est le  $p$ -laplacien, et dans ce qui suit, on va supposer :

$$\lambda > 0, \quad 1 < q < p < r < p^*, \quad \text{où } p^* = \begin{cases} \frac{Np}{N-p} & \text{if } p < N, \\ +\infty & \text{if } p \geq N. \end{cases}$$

Nous traitons un problème de limite de type Dirichlet dans le cadre de l'espace de Sobolev  $W = W_0^{1,p}(\Omega)$  muni de la norme :

$$\|u\| = \left( \int_{\Omega} |\nabla u|^p dx \right)^{1/p},$$

ce qui donne à  $W$  la structure d'un espace de Banach.

Pour les solutions de (4.1), nous étudions les points critiques de la fonctionnelle Euler-Lagrange associée  $E_{\lambda} \in C^1(W)$ , définie par :

$$E_{\lambda}(u) = \frac{1}{p}P(u) - \frac{\lambda}{q}Q(u) - \frac{1}{r}R(u), \quad (4.2)$$

où :

$$P(u) = \|u\|^p, \quad Q(u) = \int_{\Omega} |u|^q dx \quad \text{et} \quad R(u) = \int_{\Omega} |u|^r dx.$$

Ce genre de problème non-linéaire, concave-convexe, a été étudié par plusieurs auteurs [5–8, 15–18, 54]. Pour plus de détails à ce sujet, le lecteur pourra se référer à cet article [9].

Voici un résultat principal de notre travail :

- Nous déterminons une valeur  $\widehat{\lambda}$  du paramètre  $\lambda$  sous *forme variationnelle*, c'est-à-dire :

$$\widehat{\lambda} = \widehat{C}(p, q, r) \inf_{u \in W \setminus \{0\}} \mathcal{F}(u), \quad (4.3)$$

telles qu'il existe deux branches de solutions positives de (4.1) lorsque  $\lambda \in ]0, \widehat{\lambda}[$  (la fonctionnelle  $\mathcal{F}$  sera donnée ci-dessous). Notons que ces deux solutions positives sont trouvées de manière simultanée et que notre approche n'utilise pas le lemme du col (Mountain-Pass Lemma).

Par la suite, nous notons par  $\|\cdot\|_q$  et  $\|\cdot\|_r$ , les normes sur  $L^q(\Omega)$  et  $L^r(\Omega)$  respectivement. Nous introduisons comme dans [38, 96], la fonctionnelle Euler-Lagrange modifiée  $\tilde{E}_\lambda$  définie sur  $\mathbb{R} \times W$  par :

$$\tilde{E}_\lambda(t, u) := E_\lambda(tu).$$

$\partial_t \tilde{E}_\lambda(\cdot, u)$  (resp.  $\partial_{tt} \tilde{E}_\lambda(\cdot, u)$ ) où  $u$  est un élément arbitraire de  $W$ , est la dérivée (resp. la dérivée seconde) de la fonction à valeur réelle :  $t \mapsto \tilde{E}_\lambda(t, u)$ .

Notons que l'énergie  $E_\lambda$  associée au problème (4.1) n'est pas bornée inférieurement, c'est à dire que

$$\inf_{u \in W} E_\lambda(u) = -\infty.$$

Ainsi, pour mettre en évidence des solution du problème (4.1), on introduira un sous ensemble de  $W$  qui contient toutes les éventuelles solutions de (4.1) et sur lequel l'énergie  $E_\lambda$  est bornée inférieurement. On fera ainsi une minimisation sous contrainte pour montrer l'existence d'au moins deux (branches de) solutions positives du problème (4.1).

Pour calculer les deux branches de solutions, par rapport au paramètre  $\lambda$ , nous utilisons la méthode de **décomposition progressive généralisée** (PGD).

## 4.2 Résultats préliminaires

Puisque  $\tilde{E}_\lambda$  est paire en  $t$ , et comme nous sommes intéressés par les solutions non triviales de (4.1), nous limitons notre étude pour  $t > 0$  et  $u \in W \setminus \{0\}$ .

**Lemme 4.2.1.** *Pour toute  $u \in W \setminus \{0\}$ , il existe une unique  $\lambda(u) > 0$  telle que la fonction réelle  $t \mapsto \partial_t \tilde{E}_\lambda(t, u)$  s'annule exactement, pour deux valeurs positives (resp. une valeur positive) si  $0 < \lambda < \lambda(u)$  (resp.  $\lambda = \lambda(u)$ ). Cette fonction ne s'annule pas lorsque  $\lambda > \lambda(u)$ .*

**Preuve.** Soit  $u$  un élément arbitraire de  $W \setminus \{0\}$ , on a :

$$\partial_t \tilde{E}_\lambda(t, u) = t^{q-1} \tilde{F}_\lambda(t, u), \text{ où } \tilde{F}_\lambda(t, u) = t^{p-q} P(u) - \lambda Q(u) - t^{r-q} R(u).$$

alors :

$$\partial_{tt} \tilde{E}_\lambda(t, u) = (q-1)t^{q-2} \tilde{F}_\lambda(t, u) + t^{q-1} \partial_t \tilde{F}_\lambda(t, u),$$

avec :

$$\partial_t \tilde{F}_\lambda(t, u) = t^{p-q-1} \{ (p-q)P(u) - (r-q)t^{r-p} R(u) \}.$$

Il est clair que la fonction réelle  $t \mapsto \tilde{F}_\lambda(t, u)$  est croissante sur  $]0, t(u)[$ , décroissante sur  $]t(u), +\infty[$ , et elle atteint son unique maximum pour  $t = t(u)$ , où :

$$t(u) = \left( \frac{p-q}{r-q} \frac{P(u)}{R(u)} \right)^{\frac{1}{r-p}}. \quad (4.4)$$



Donc, la fonction  $t \mapsto \tilde{F}_\lambda(t, u)$  s'annule dans deux points (resp. un seul point) si  $\tilde{F}_\lambda(t(u), u) > 0$  (resp. si  $\tilde{F}_\lambda(t(u), u) = 0$ ) et elle ne s'annule pas si  $\tilde{F}_\lambda(t(u), u) < 0$ . D'un autre côté, un calcul direct donne :

$$\tilde{F}_\lambda(t(u), u) = \frac{r-p}{p-q} \left( \frac{p-q}{r-q} \frac{P(u)}{R(u)} \right)^{\frac{r-q}{r-p}} R(u) - \lambda Q(u).$$

Pareillement,  $\tilde{F}_\lambda(t(u), u) > 0$  (resp.  $\tilde{F}_\lambda(t(u), u) < 0$ ) si  $\lambda < \lambda(u)$  (resp.  $\lambda > \lambda(u)$ ) et  $\tilde{F}_{\lambda(u)}(t(u), u) = 0$ , où :

$$\lambda(u) = \widehat{C} \frac{P^{\frac{r-q}{r-p}}(u)}{Q(u) R^{\frac{p-q}{r-p}}(u)}, \quad (4.5)$$

avec :

$$\widehat{C} = \frac{r-p}{p-q} \left( \frac{p-q}{r-q} \right)^{\frac{r-q}{r-p}}.$$

Donc, si  $\lambda \in ]0, \lambda(u)[$ , la fonction réelle  $t \mapsto \partial_t \tilde{E}_\lambda(t, u)$  s'annule dans deux points, notés par :  $\underline{t}(u, \lambda)$  et  $\bar{t}(u, \lambda)$ , vérifiant  $0 < \underline{t}(u, \lambda) < t(u) < \bar{t}(u, \lambda)$ .

Puisque,  $\tilde{F}_\lambda(\underline{t}(u, \lambda), u) = \tilde{F}_\lambda(\bar{t}(u, \lambda), u) = 0$ ,  $\partial_t \tilde{F}_\lambda(t, u) > 0$  pour  $t < t(u)$  et  $\partial_t \tilde{F}_\lambda(t, u) < 0$  pour  $t > t(u)$ , il s'en suit :

$$\partial_{tt} \tilde{E}_\lambda(\underline{t}(u, \lambda), u) > 0 \quad \text{et} \quad \partial_{tt} \tilde{E}_\lambda(\bar{t}(u, \lambda), u) < 0.$$

Ceci implique que la fonction réelle  $t \mapsto \tilde{E}_\lambda(t, u)$ ,  $t > 0$ , atteint son unique minimum local en  $t = \underline{t}(u, \lambda)$  et son maximum global en  $t = \bar{t}(u, \lambda)$ .  $\square$

Précisons que pour toute  $u \in W \setminus \{0\}$  et  $\lambda \in ]0, \lambda(u)[$ , nous avons  $\underline{t}(u, \lambda)u$  et  $\bar{t}(u, \lambda)u$  appartiennent à la variété de Nehari [75, 96] définie par :

$$\mathcal{N} := \{u \in W \setminus \{0\} : E'_\lambda(u)u = 0\}.$$

Maintenant, nous introduisons :

$$\widehat{\lambda} := \inf_{u \in W \setminus \{0\}} \lambda(u). \quad (4.6)$$

Si  $c_q(\Omega)$  (resp.  $c_r(\Omega)$ ) signifie la constante de Sobolev de l'injection  $W \subset L^q(\Omega)$  (resp.  $W \subset L^r(\Omega)$ ), alors :

$$\widehat{\lambda} \geq \widehat{C} [c_q(\Omega)]^{q/p} [c_r(\Omega)]^{\frac{r(p-q)}{p(r-p)}} > 0.$$

Comme  $\partial_t \tilde{E}_\lambda(\underline{t}(u, \lambda), u) = 0$  (resp.  $\partial_t \tilde{E}_\lambda(\bar{t}(u, \lambda), u) = 0$ ) pour toute  $u \in W \setminus \{0\}$ , alors la fonctionnelle  $u \mapsto \tilde{E}_\lambda(\underline{t}(u, \lambda), u)$  (resp.  $u \mapsto \tilde{E}_\lambda(\bar{t}(u, \lambda), u)$ ) est bornée inférieure sur  $W \setminus \{0\}$ . Alors, pour toute  $\lambda \in ]0, \widehat{\lambda}[$ , nous définissons :

$$\underline{\alpha}(\lambda) = \inf_{u \in W \setminus \{0\}} \tilde{E}_\lambda(\underline{t}(u, \lambda), u), \quad (4.7)$$

$$\overline{\alpha}(\lambda) = \inf_{u \in W \setminus \{0\}} \tilde{E}_\lambda(\bar{t}(u, \lambda), u). \quad (4.8)$$

**Lemme 4.2.2.** Soit  $(u_n) \subset W \setminus \{0\}$  une suite minimisante de (4.7) (resp. de (4.8)) et  $\underline{U}_n := \underline{t}(u_n, \lambda)u_n$  (resp.  $\overline{U}_n := \overline{t}(u_n, \lambda)u_n$ ). Alors :

- (i)  $\limsup_{n \rightarrow +\infty} \|\underline{U}_n\| < +\infty$  (resp.  $\limsup_{n \rightarrow +\infty} \|\overline{U}_n\| < +\infty$ ),
- (ii)  $\liminf_{n \rightarrow +\infty} \|\underline{U}_n\| > 0$  (resp.  $\liminf_{n \rightarrow +\infty} \|\overline{U}_n\| > 0$ ).

**Preuve.** (i) Soit  $(u_n) \subset W \setminus \{0\}$  une suite minimisante de (4.7). Comme  $\partial_t \tilde{E}_\lambda(\underline{t}(u_n, \lambda), u_n) = 0$ , alors :

$$\|\underline{U}_n\|^p = \lambda \|\underline{U}_n\|_q^q + \|\underline{U}_n\|_r^r. \quad (4.9)$$

De la même façon, puisque  $\partial_{tt} \tilde{E}_\lambda(\underline{t}(u_n, \lambda), u_n) > 0$ , alors :

$$(p-1)\|\underline{U}_n\|^p - \lambda(q-1)\|\underline{U}_n\|_q^q - (r-1)\|\underline{U}_n\|_r^r > 0. \quad (4.10)$$

En combinant (4.9) et (4.10), nous trouvons  $E_\lambda(\underline{U}_n) < 0$ , pour tout  $n$ .

Supposons qu'il existe une sous-suite de  $(\underline{U}_n)$ , que l'on note encore  $(\underline{U}_n)$  telle que  $\lim_{n \rightarrow +\infty} \|\underline{U}_n\| = +\infty$ . Il est connu qu'il existe une certaine constante  $C$ , telle que  $\|\underline{U}_n\|_q \leq C\|\underline{U}_n\|_r$ , pour tout  $n$ , alors  $\lim_{n \rightarrow +\infty} \|\underline{U}_n\|_r = +\infty$ . En utilisant le fait que  $0 < q < r$ , nous obtenons  $\|\underline{U}_n\|_q^q = o_n(\|\underline{U}_n\|_r^r)$ , et par conséquent :

$$\|\underline{U}_n\|^p = \|\underline{U}_n\|_r^r(1 + o_n(1)).$$

Ainsi,

$$E_\lambda(\underline{U}_n) = \|\underline{U}_n\|_r^r \left( \frac{1}{p} - \frac{1}{r} + o_n(1) \right),$$

ce qui implique que  $E_\lambda(\underline{U}_n)$  tend vers  $+\infty$  lorsque  $n$  tend vers  $+\infty$ . Ceci est impossible. On conclut que  $\limsup_{n \rightarrow +\infty} \|\underline{U}_n\| < +\infty$ . Les mêmes arguments, avec une suite minimisante  $(u_n)$  de (4.8), prouvent que  $\limsup_{n \rightarrow +\infty} \|\overline{U}_n\| < +\infty$ .

(ii) Soit  $(u_n) \subset W \setminus \{0\}$  une suite minimisante de (4.7) et supposons qu'il existe une sous-suite minimisante de  $(\underline{U}_n)$ , que l'on note encore  $(\underline{U}_n)$  telle que  $\lim_{n \rightarrow +\infty} \|\underline{U}_n\| = 0$ . Ceci nous donne  $\lim_{n \rightarrow +\infty} E_\lambda(\underline{U}_n) = 0$ , c'est-à-dire  $\underline{\alpha}(\lambda) = 0$ . Ce qui est impossible, car  $\tilde{E}_\lambda(\underline{t}(u_n, \lambda), u_n) < 0$  pour tout  $n$ .

Soit  $(u_n) \subset W \setminus \{0\}$  une suite minimisante de (4.8). Comme  $\partial_t \tilde{E}_\lambda(\overline{t}(u_n, \lambda), u_n) = 0$  et  $\partial_{tt} \tilde{E}_\lambda(\overline{t}(u_n, \lambda), u_n) < 0$ , ceci implique :

$$\begin{cases} \|\overline{U}_n\|^p - \lambda \|\overline{U}_n\|_q^q - \|\overline{U}_n\|_r^r = 0, \\ (p-1)\|\overline{U}_n\|^p - \lambda(q-1)\|\overline{U}_n\|_q^q - (r-1)\|\overline{U}_n\|_r^r < 0. \end{cases}$$

En combinant les deux dernières inégalités, on obtient, pour tout  $n$  :

$$(p-q)\|\overline{U}_n\|^p < (r-q)\|\overline{U}_n\|_r^r \leq C'\|\overline{U}_n\|^r,$$

par l'injection continue  $W \subset L^r(\Omega)$ . Alors  $(p - q) \leq C' \|\bar{U}_n\|^{r-p}$ . À ce stade, supposons qu'il existe une sous-suite de  $(\bar{U}_n)$ , encore notée par  $(\bar{U}_n)$  telle que  $\lim_{n \rightarrow +\infty} \|\bar{U}_n\| = 0$ . Ce qui implique que  $p - q \leq 0$ . Ceci est impossible.  $\square$

**Remarque 4.2.3.** Puisque, pour tout nombre réel  $\gamma > 0$ , on a :

$$\begin{aligned} \tilde{E}_\lambda \left( \gamma t, \frac{u}{\gamma} \right) &= \tilde{E}_\lambda(t, u), \\ \partial_t \tilde{E}_\lambda \left( \gamma t, \frac{u}{\gamma} \right) &= \frac{1}{\gamma} \partial_t \tilde{E}_\lambda(t, u), \\ \partial_{tt} \tilde{E}_\lambda \left( \gamma t, \frac{u}{\gamma} \right) &= \frac{1}{\gamma^2} \partial_{tt} \tilde{E}_\lambda(t, u), \end{aligned}$$

ceci implique que :

$$\underline{t}(u, \lambda) = \frac{1}{\gamma} \underline{t} \left( \frac{u}{\gamma}, \lambda \right), \quad (4.11)$$

$$\bar{t}(u, \lambda) = \frac{1}{\gamma} \bar{t} \left( \frac{u}{\gamma}, \lambda \right). \quad (4.12)$$

Par suite, on conclut que :

$$\underline{\alpha}(\lambda) = \inf_{u \in \mathbb{S}} \tilde{E}_\lambda(\underline{t}(u, \lambda), u), \quad (4.13)$$

$$\bar{\alpha}(\lambda) = \inf_{u \in \mathbb{S}} \tilde{E}_\lambda(\bar{t}(u, \lambda), u), \quad (4.14)$$

où  $\mathbb{S}$  est la sphère unité de  $W$ .

**Lemme 4.2.4.** Soit  $(u_n) \subset \mathbb{S}$  une suite minimisante de (4.13) (resp. de (4.14)). Alors,  $(\underline{U}_n) := (\underline{t}(u_n, \lambda)u_n)$  (resp.  $(\bar{U}_n) := (\bar{t}(u_n, \lambda)u_n)$ ) sont des suites de Palais-Smale pour la fonctionnelle  $E_\lambda$ .

**Preuve.** On va démontrer ce lemme juste pour la suite  $(\underline{U}_n)$ . On peut faire la démonstration de  $(\bar{U}_n)$  de la même façon.

D'abord, d'après le lemme précédent, il est clair que  $(\underline{U}_n)$  est bornée dans  $W$ . D'autre part, notons que pour toute  $u \in W \setminus \{0\}$  et  $\lambda \in ]0, \widehat{\lambda}[$ , on a  $\partial_t \tilde{E}_\lambda(\underline{t}(u, \lambda), u) = 0$  et  $\partial_{tt} \tilde{E}_\lambda(\underline{t}(u, \lambda), u) \neq 0$ . Puisque  $\tilde{E}$  est de classe  $C^1$  par rapport à  $u$ , en utilisant le théorème de la fonction implicite, on trouve que  $\underline{t}(u, \lambda)$  est aussi de classe  $C^1$  par rapport à  $u$ . Soit  $\underline{\mathcal{E}}_\lambda$  une fonctionnelle de classe  $C^1$  définie sur  $\mathbb{S}$  par :

$$\underline{\mathcal{E}}_\lambda(u) = \tilde{E}_\lambda(\underline{t}(u, \lambda), u) \equiv E_\lambda(\underline{t}(u, \lambda)u).$$

Alors :

$$\underline{\alpha}(\lambda) = \inf_{u \in \mathbb{S}} \underline{\mathcal{E}}_\lambda(u) \quad \text{et} \quad \lim_{n \rightarrow +\infty} \underline{\mathcal{E}}_\lambda(u_n) = \underline{\alpha}(\lambda).$$

En utilisant le principe variationnel d'Ekeland sur la variété complète  $(\mathbb{S}, \|\cdot\|)$  de la fonctionnelle  $\underline{\mathcal{E}}_\lambda$ , on conclut que :

$$|\underline{\mathcal{E}}'_\lambda(u_n)(\varphi_n)| \leq \frac{1}{n} \|\varphi_n\|, \text{ quelque soit } \varphi_n \in T_{u_n}\mathbb{S},$$

où  $T_{u_n}\mathbb{S}$  est l'espace tangent sur  $\mathbb{S}$  dans le point  $u_n$ . De plus, quelque soit  $\varphi_n \in T_{u_n}\mathbb{S}$ , on a :

$$\begin{aligned} \underline{\mathcal{E}}'_\lambda(u_n)(\varphi_n) &= \partial_t \tilde{E}_\lambda(\underline{t}(u_n, \lambda), u_n) \underline{t}'(u_n, \lambda)(\varphi_n) + \partial_u \tilde{E}_\lambda(\underline{t}(u_n, \lambda), u_n)(\varphi_n), \\ &= \partial_u \tilde{E}_\lambda(\underline{t}(u_n, \lambda), u_n)(\varphi_n), \end{aligned}$$

puisque  $\partial_t \tilde{E}_\lambda(\underline{t}(u_n, \lambda), u_n) \equiv 0$ , où  $\underline{t}'(u_n, \lambda)$  signifie la dérivée de  $\underline{t}(\cdot, \lambda)$  par rapport à sa première variable, au point  $(u_n, \lambda)$ .

En outre, soit :

$$\begin{aligned} \pi : W \setminus \{0\} &\longrightarrow \mathbb{R} \times \mathbb{S} \\ u &\longmapsto \left( \|u\|, \frac{u}{\|u\|} \right) := (\pi_1(u), \pi_2(u)). \end{aligned}$$

Par l'inégalité de Hölder, on trouve pour toute  $(u, \varphi) \in (W \setminus \{0\}) \times W$  :

$$\begin{cases} |\pi'_1(u)(\varphi)| &\leq \|\varphi\|, \\ \|\pi'_2(u)(\varphi)\| &\leq 2 \frac{\|\varphi\|}{\|u\|}. \end{cases}$$

Par le lemme 4.2.2, il existe une constante positive  $C$  telle que :

$$\underline{t}(u_n, \lambda) \geq C, \quad \forall n \in \mathbb{N}.$$

Alors, quelque soit  $\varphi \in W$ , ils existent  $\varphi_n^1 \in \mathbb{R}$  et  $\varphi_n^2 \in T_{u_n}\mathbb{S}$  telles que  $|\varphi_n^1| \leq \|\varphi\|$ ,  $\|\varphi_n^2\| \leq \frac{2}{C} \|\varphi\|$  et :

$$\begin{aligned} E'_\lambda(\underline{t}(u_n, \lambda)u_n)(\varphi) &= \partial_t \tilde{E}_\lambda(\underline{t}(u_n, \lambda), u_n)(\varphi_n^1) + \partial_u \tilde{E}_\lambda(\underline{t}(u_n, \lambda), u_n)(\varphi_n^2), \\ &= \partial_u \tilde{E}_\lambda(\underline{t}(u_n, \lambda), u_n)(\varphi_n^2), \\ &= \underline{\mathcal{E}}'_\lambda(u_n)(\varphi_n^2). \end{aligned}$$

D'où :

$$\begin{aligned} E'_\lambda(\underline{t}(u_n, \lambda)u_n)(\varphi) &\leq \frac{1}{n} \|\varphi_n^2\| \\ &\leq \frac{2}{nC} \|\varphi\|. \end{aligned}$$

On conclut facilement :

$$\lim_{n \rightarrow \infty} \|E'_\lambda(\underline{U}_n)\|_* = 0.$$

□

**Remarque 4.2.5.** *Jusque là, toutes les suites minimisantes considérées ne sont pas positives. Notons que, pour toute  $u \in W \setminus \{0\}$  et  $0 < \lambda < \widehat{\lambda}$ , on a  $\tilde{E}_\lambda(t, |u|) = \tilde{E}_\lambda(t, u)$ ,  $\underline{t}(|u|, \lambda) = \underline{t}(u, \lambda)$  et  $\bar{t}(|u|, \lambda) = \bar{t}(u, \lambda)$ . Alors, chaque suite minimisante  $(u_n) \subset \mathbb{S}$  de (4.13) ou bien (4.14) peut être considérée comme une suite des fonctions positives.*

Dorénavant, nous allons supposer les suites  $\underline{U}_n$  et  $\overline{U}_n$ , définies au lemme 4.2.4, positives.

### 4.3 Résultats d'existence

**Théorème 4.3.1.** *Soient  $1 < q < p < r < p^*$  et  $\lambda \in ]0, \widehat{\lambda}[$ . Alors le problème (4.1) admet, au moins, deux solutions.*

**Preuve.** D'après le lemme 4.2.4, on a :  $E_\lambda(\underline{U}_n)$  converge vers  $\underline{\alpha}(\lambda)$ ,  $\|E'_\lambda(\underline{U}_n)\|_*$  converge vers 0 lorsque  $n$  tend vers  $+\infty$  et que  $(\underline{U}_n)$  est bornée dans  $W$ . Si nécessaire, passons à une sous-suite, et on obtient :

$$\begin{aligned} \underline{U}_n &\rightharpoonup \underline{U} \quad \text{dans } W, \\ \underline{U}_n &\rightarrow \underline{U} \quad \text{dans } L^r(\Omega), \quad (\text{encore dans } L^q(\Omega)), \\ \underline{U}_n &\rightarrow \underline{U} \quad \text{p.p. } \Omega. \end{aligned}$$

Soit  $v_n = \underline{U}_n - \underline{U}$ , alors en utilisant le lemme de Brezis-Lieb [19], on trouve :

$$\begin{aligned} \|v_n\| &= \|\underline{U}_n\| - \|\underline{U}\| + o_n(1), \\ \|v_n\|_q &= \|\underline{U}_n\|_q - \|\underline{U}\|_q + o_n(1), \\ \|v_n\|_r &= \|\underline{U}_n\|_r - \|\underline{U}\|_r + o_n(1). \end{aligned}$$

Ce qui nous donne :

$$\begin{aligned} E_\lambda(v_n) &= E_\lambda(\underline{U}_n) - E_\lambda(\underline{U}) + o_n(1), \\ E'_\lambda(v_n) &= E'_\lambda(\underline{U}_n) - E'_\lambda(\underline{U}) + o_n(1), \end{aligned}$$

et par conséquent,  $E'_\lambda(v_n)v_n \rightarrow 0$  lorsque  $n \rightarrow +\infty$ . Ceci implique :

$$\|\underline{v}_n\|^p = \lambda \|\underline{v}_n\|_q^q + \|\underline{v}_n\|_r^r + o_n(1).$$

Donc,  $\|v_n\|^p \rightarrow 0$  lorsque  $n \rightarrow +\infty$ . On peut utiliser le même argument pour  $\overline{U}_n$ . On sait que pour toute suite minimisante  $(u_n)$  de (4.13), il existe une sous-suite notée par  $(u_n)$  telle que :

$$\underline{U}_n := \underline{t}(u_n, \lambda)u_n \longrightarrow \underline{U} \quad \text{dans } W.$$

De plus, d'après le lemme 4.2.2 et le théorème 4.3.1, il existe  $\underline{t} \in ]0, +\infty[$  tel que :

$$\begin{cases} \underline{t}(u_n, \lambda) & \longrightarrow \underline{t} \text{ dans } \mathbb{R}, \\ u_n & \longrightarrow \underline{u} = \underline{U}/\underline{t} \text{ dans } W, \end{cases}$$

avec  $\underline{u} = \underline{U}/\underline{t} \in \mathbb{S}$ , (c'est-à-dire  $\underline{t} = \|\underline{U}\|$ ) et  $\underline{t} = \underline{t}(\underline{u}, \lambda)$ . De la même façon, pour toute suite minimisante  $(w_n)$  de (4.14), si c'est nécessaire, passons à une sous-suite, il existe  $\bar{t} \in ]0, +\infty[$  tel que :

$$\begin{cases} \bar{t}(w_n, \lambda) & \longrightarrow \bar{t} \text{ dans } \mathbb{R}, \\ w_n & \longrightarrow \bar{u} = \bar{U}/\bar{t} \text{ dans } W, \end{cases}$$

avec  $\bar{u} = \bar{U}/\bar{t} \in \mathbb{S}$ , (c'est-à-dire  $\bar{t} = \|\bar{U}\|$ ) et  $\bar{t} = \bar{t}(\bar{u}, \lambda)$ . À ce stade, nous voyons facilement que  $\underline{U} \neq \bar{U}$ . En effet, puisque  $\partial_{tt}\tilde{E}_\lambda(\underline{t}(\underline{u}, \lambda), \underline{u}) > 0$  et  $\partial_{tt}\tilde{E}_\lambda(\bar{t}(\bar{u}, \lambda), \bar{u}) < 0$ , alors  $\partial_{tt}\tilde{E}_\lambda(\underline{t}, \underline{U}/\underline{t}) > 0$  et  $\partial_{tt}\tilde{E}_\lambda(\bar{t}, \bar{U}/\bar{t}) < 0$ . Ce qui termine la preuve.  $\square$

## 4.4 Decomposition Progressive Généralisée (PGD)

Dans cette section, on suppose que  $\Omega$  est un ouvert borné de  $\mathbb{R}^N$ ,  $N \geq 2$ , de la forme

$$\Omega = \prod_{k=1}^N \Omega_k, \quad k = 1, 2, \dots, N$$

où les  $\Omega_k$  sont des ouverts bornés de  $\mathbb{R}$ . L'espace fonctionnel associé au problème (4.1) est  $W = W_0^{1,p}(\Omega)$  muni de sa norme classique  $\|\cdot\|$  et le caractère tensoriel de  $W$  s'écrit :

$$W = \overline{\bigotimes_{k=1}^N W_0^{1,p}(\Omega_k)}^{\|\cdot\|},$$

où  $\bigotimes_{k=1}^N W_0^{1,p}(\Omega_k)$  désigne le produit tensoriel algébrique des espaces de Sobolev (en dimension 1)  $W_0^{1,p}(\Omega_k)$  (voir Chapitre 1). Dans la suite, l'ensemble des tenseurs élémentaires de  $\bigotimes_{k=1}^N W_0^{1,p}(\Omega_k)$  sera noté par  $\mathcal{R}_1(W)$ . Ainsi l'espace engendré par  $\mathcal{R}_1(W)$  est dense dans  $W$  pour la norme  $\|\cdot\|$ .

Pour chaque  $u \in W$ , on considère le problème

$$\inf_{z \in \mathcal{R}_1(W), z \neq -u} E_\lambda(\underline{t}(u+z, \lambda)(u+z)) \quad (4.15)$$

Montrer que cet infimum est atteint est une question complexe due au fait que la variété de Nehari n'est pas faiblement fermée dans  $W$ , de plus, la fonctionnelle définie sur  $W \setminus \{0\}$  par  $v \mapsto E_\lambda(\underline{t}(v, \lambda)v)$  n'a aucune propriété de convexité. Dans la suite, on admettra que l'infimum du problème (4.15) est atteint pour tout  $u \in W$ .

Dans la suite, nous considérons l'approche suivante de la PGD :

$$(M) \quad \begin{cases} (i) & \text{Initialisation : } u_0 = 0 \quad \text{dans } W \\ (ii) & \text{Choisir } z_m := \arg \min_{z \in \mathcal{R}_1(W)} E_\lambda(\underline{t}(u_{m-1} + z, \lambda)(u_{m-1} + z)). \\ (iii) & u_m := u_{m-1} + z_m, \quad \forall m \geq 1. \end{cases}$$

Il s'agit de la méthode la plus utilisée en PGD, grâce à sa simplicité d'implémentation. Nous montrerons que toute suite PGD de type (M) est faiblement convergente dans  $W_{\|\cdot\|}$ , pour  $p = 2$ .

**Théorème 4.4.1.** Soit  $\Omega = \prod_{k=1}^N \Omega_k$  un domaine borné de  $\mathbb{R}^N$ , où  $(\Omega_k)_{1 \leq k \leq N}$  désignent des intervalles ouverts bornés de  $\mathbb{R}$ . Soit  $E_\lambda$  la fonctionnelle définie sur  $W$  par (4.2). Supposons qu'il existe  $m \in \mathbb{N}$  tel que :  $E(\underline{t}(u_m, \lambda)u_m) = E(\underline{t}(u_{m+1}, \lambda)u_{m+1})$ , alors  $\underline{t}(u_m, \lambda)u_m$  est solution de (4.1).

En procédant de la même manière avec une suite PGD du type  $(\bar{t}(u_m, \lambda)u_m)_{m \in \mathbb{N}}$ , on aboutit en cas de convergence à une seconde solution de (4.1) de la forme  $(\bar{t}(\bar{u}, \lambda)\bar{u}) \neq 0$ .

**Preuve.** Dans le cas où  $p = 2$ , l'espace  $W = H_0^1(\Omega)$ . Soit une suite  $(u_m)_{m \in \mathbb{N}} \subset W$  de type PGD, définie par (M). Il est clair que la suite réelle  $(E_\lambda(\underline{t}(u_m, \lambda)u_m))_{m \in \mathbb{N}}$  est décroissante, minorée et par conséquent convergente dans  $\mathbb{R}$ , on notera  $\ell$  sa limite. D'autre part, on sait d'après ce qui précède que la suite  $(\underline{t}(u_m, \lambda)u_m)_{m \in \mathbb{N}}$  est bornée dans  $W$ .

Notons par :  $\mathcal{N} := \underline{\mathcal{N}} \cup \overline{\mathcal{N}}$  et

$$\begin{aligned} F : W \setminus \{0\} &\longrightarrow \underline{\mathcal{N}} \\ u &\longmapsto \underline{t}(u, \lambda)u \end{aligned}$$

où  $\underline{\mathcal{N}} := \{\underline{t}(u, \lambda)u \mid u \in W \setminus \{0\}\}$ .

$F$  est de classe  $C^1$  car  $\underline{t}(u, \lambda)$  l'est. Notons par :

$$\begin{aligned} E : W \setminus \{0\} &\longrightarrow \mathbb{R} \\ u &\longmapsto E \circ F(u) = E(\underline{t}(u, \lambda)u) \end{aligned}$$

Soit  $m \in \mathbb{N}$  tel que :  $E(\underline{t}(u_m, \lambda)u_m) = E(\underline{t}(u_{m+1}, \lambda)u_{m+1})$ , c'est-à-dire  $E \circ F(u_m) \leq E \circ F(u_m + z) \quad \forall z \in \mathcal{R}_1$  donc  $E'(F(u)) \cdot F'(u) = 0 \quad \forall z \in \mathcal{R}_1$ , or  $F'(u) \cdot z = (\underline{t}'(u) \cdot z)u + \underline{t}(u) \cdot z$ .

D'où :

$$E'(\underline{t}(u, \lambda)u)(\underline{t}'(u) \cdot z)u + E'(\underline{t}(u, \lambda)u) \cdot z = 0 \quad \forall z \in \mathcal{R}_1$$

ceci implique :

$$E'(\underline{t}(u, \lambda)u) \cdot z = 0, \quad \forall z \in \mathcal{R}_1.$$

Par densité de  $\mathcal{R}_1$  dans  $W$  on trouve :

$$E'(\underline{t}(u, \lambda)u) \cdot z = 0, \quad \forall z \in W.$$

ce qui termine la preuve. De même pour la suite  $(\bar{t}(u_m, \lambda)u_m)_{m \in \mathbb{N}}$ , d'où le résultat.  $\square$

## 4.5 Résultats numériques

Dans ce qui suit, on va mettre en évidence deux suites PGD qui convergent vers deux solutions positives de la forme  $(\underline{t}(\underline{u}, \lambda)\underline{u})$  et  $(\bar{t}(\bar{u}, \lambda)\bar{u})$  respectivement.

Pour la première solution, notons que toute solution du problème

$$\inf_{u \in W \setminus \{0\}} E_\lambda(\underline{t}(u, \lambda)u)$$

est forcément solution de l'équation aux dérivées partielles :

$$\underline{t}(u, \lambda)^p \Delta_p u = \lambda \underline{t}(u, \lambda)^q |u|^{q-2} u + \underline{t}(u, \lambda)^r |u|^{r-2} u \quad \text{dans } \Omega,$$

avec des conditions aux limites de type Dirichlet homogènes. Les mêmes arguments sont valables pour la solution du problème

$$\inf_{u \in W \setminus \{0\}} E_\lambda(\bar{t}(u, \lambda)u).$$

Ce qui fournit deux solutions positives distinctes du problème (4.1).

Toutes les simulations numériques ont été réalisées sur le domaine  $\Omega = ]0, 1[ \times ]0, 1[$  avec des conditions aux limites de Dirichlet homogènes et  $\lambda = 0.01$ ,  $q = 1.2$  et  $r = 2.9$ . La discrétisation a été faite via la méthode des volumes finis et le nombre de noeuds dans la direction  $x$  est le même que celui dans la direction  $y$  et il vaut 51.

### • Méthode classique :

Tant que  $\|U^{(k)} - U^{(k-1)}\|_{L^2} \leq \varepsilon_{\text{picard}}$  ( $k$  désigne l'itération de Picard)

– on résout

$$-\Delta U^{(k)} = \lambda |U^{(k-1)}|^{q-2} U^{(k-1)} + |U^{(k-1)}|^{r-2} U^{(k-1)},$$

où  $U(x, y)$  désigne  $\underline{t}(u, \lambda)u$  ou bien  $\bar{t}(u, \lambda)u$ , soit par la méthode du bi-gradient conjugué (appelée modèle complet), soit par PGD en recherchant la solution sous la forme séparée

$$U^{(k)}(x, y) = \sum_{i=1}^{N_{pgd}} F^{(k)}(x) G^{(k)}(y).$$



• Méthode proposée :

- Tant que  $\|U^{(k)} - U^{(k-1)}\|_{L^2} \leq \varepsilon_{picard}$  ( $k$  désigne l'itération de Picard)
- Calcul des valeurs des polynômes  $P$ ,  $Q$  et  $R$  en  $U^{(k-1)}$  :

$$\begin{aligned} P(U^{(k-1)}) &= \int_{\Omega} |\nabla U^{(k-1)}|^2 dx dy, \\ Q(U^{(k-1)}) &= \int_{\Omega} |U^{(k-1)}|^q dx dy, \\ R(U^{(k-1)}) &= \int_{\Omega} |U^{(k-1)}|^r dx dy. \end{aligned}$$

- Recherche de la racine  $t_0$  de la fonction

$$F(t) = \frac{t^2}{2}P(U^{(k-1)}) - \lambda \frac{t^q}{q}Q(U^{(k-1)}) - \frac{t^r}{r}R(U^{(k-1)})$$

- Calcul du second membre  $SM^{(k-1)}$  :

$$SM^{(k-1)} = \lambda t_0^{q-2} |U^{(k-1)}|^{q-2} U^{(k-1)} + t_0^{r-2} |U^{(k-1)}|^{r-2} U^{(k-1)}$$

- on résout

$$-\Delta U^{(k)} = SM^{(k-1)}$$

soit par la méthode du bi-gradient conjugué (appelée modèle complet),  
soit par PGD en recherchant la solution sous la forme séparée

$$U^{(k)}(x, y) = \sum_{i=1}^{N_{pgd}} F^{(k)}(x) G^{(k)}(y)$$

• Remarques :

- 1) Pour obtenir la "solution"
  - associée à  $t_{min}$  c'est-à-dire là où la fonction  $F$  est convexe, la racine de  $F(t)$  est cherchée sur l'intervalle  $[0 : 2]$  (en utilisant Newton)
  - associée à  $t_{max}$  c'est à dire là où la fonction  $F$  est concave, la racine de  $F(t)$  est cherchée sur l'intervalle  $[0 : 1000]$  (en utilisant Newton)
- 2) Pour les calculs  $\varepsilon_{picard} = 10^{-13}$
- 3) Avec la méthode classique, seule la solution correspondant à  $t_{min}$  est obtenue. Dans ce cas que ce soit en complet ou en PGD l'algorithme converge en 9 itérations
- 4) Avec la méthode proposée, que ce soit en complet ou en PGD, l'algorithme de Picard converge en 4 itérations pour la solution associée à  $t_{min}$  et 18 itérations pour la solution associée à  $t_{max}$ .

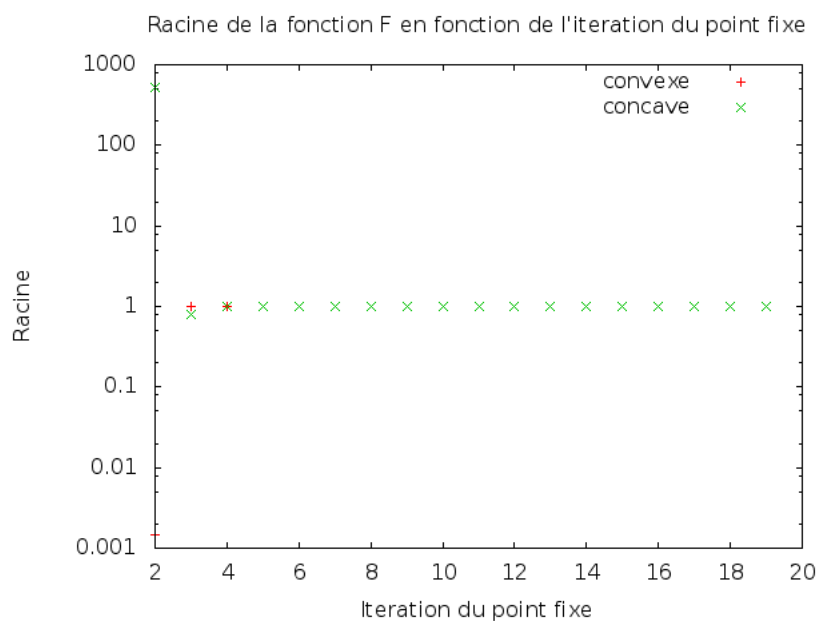


FIGURE 4.1 – Evolution de la racine de la fonction  $F(t)$  en fonction de l'itération de l'algorithme de Picard.

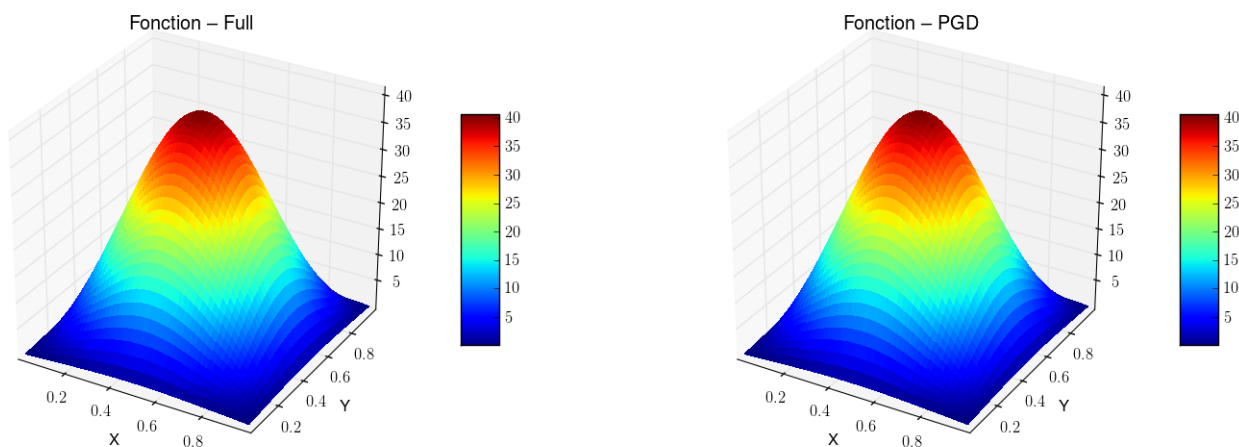


FIGURE 4.2 – Solutions obtenues par résolution complète et par résolution PGD (solutions obtenues pour  $\bar{t}_{max}$  c'est à dire là où la fonction  $F$  est concave) – représentation 3D

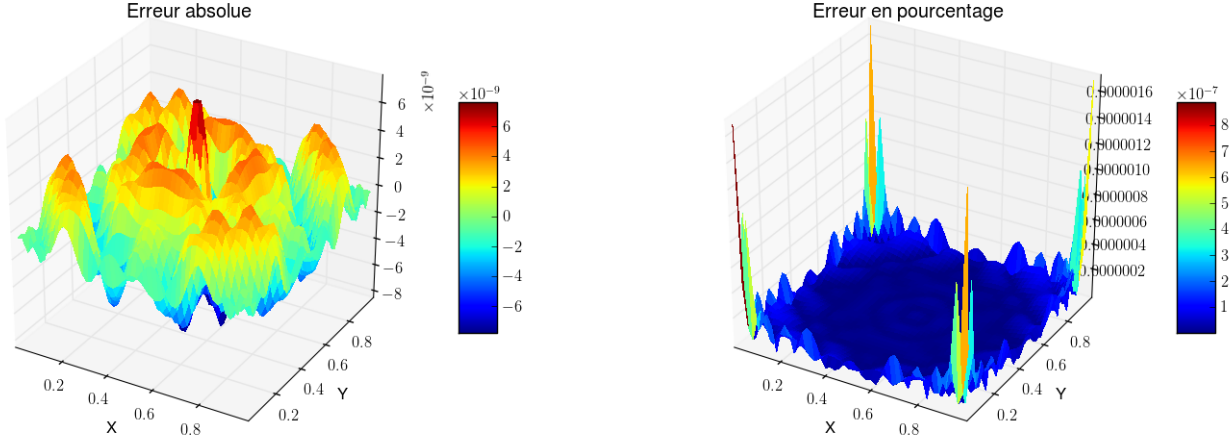


FIGURE 4.3 – Comparaison des solutions obtenues par résolution complète et par résolution PGD (solution obtenue pour  $\bar{t}_{max}$  c'est à dire là où la fonction F est concave)– représentation 3D

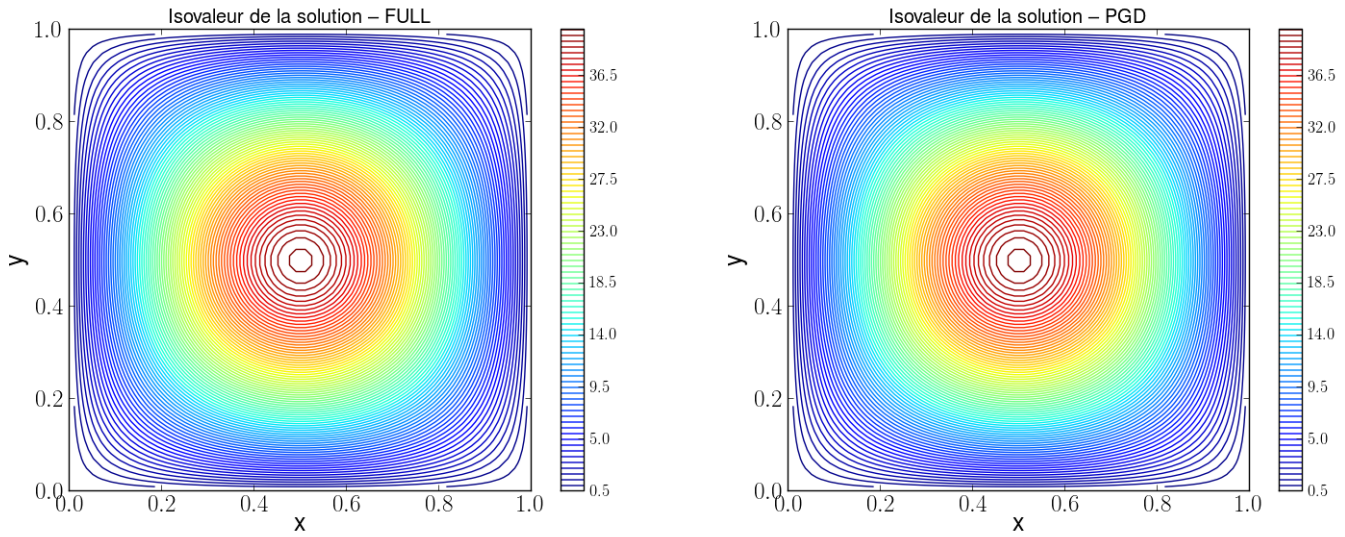


FIGURE 4.4 – Solutions obtenues par résolution complète et par résolution PGD (solutions obtenues pour  $\bar{t}_{max}$  c'est à dire là où la fonction F est concave) – représentation 2D

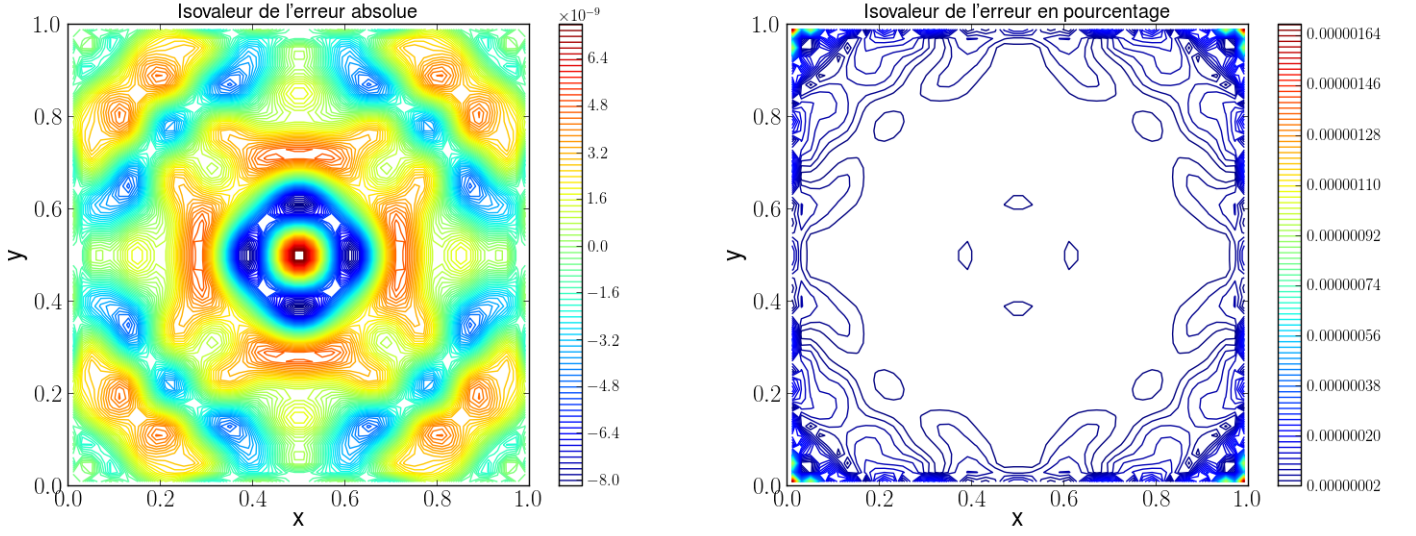


FIGURE 4.5 – Comparaison des solutions obtenues par résolution complète et par résolution PGD (solution obtenue pour  $\bar{t}_{max}$  c'est à dire là où la fonction  $F$  est concave)– représentation 2D

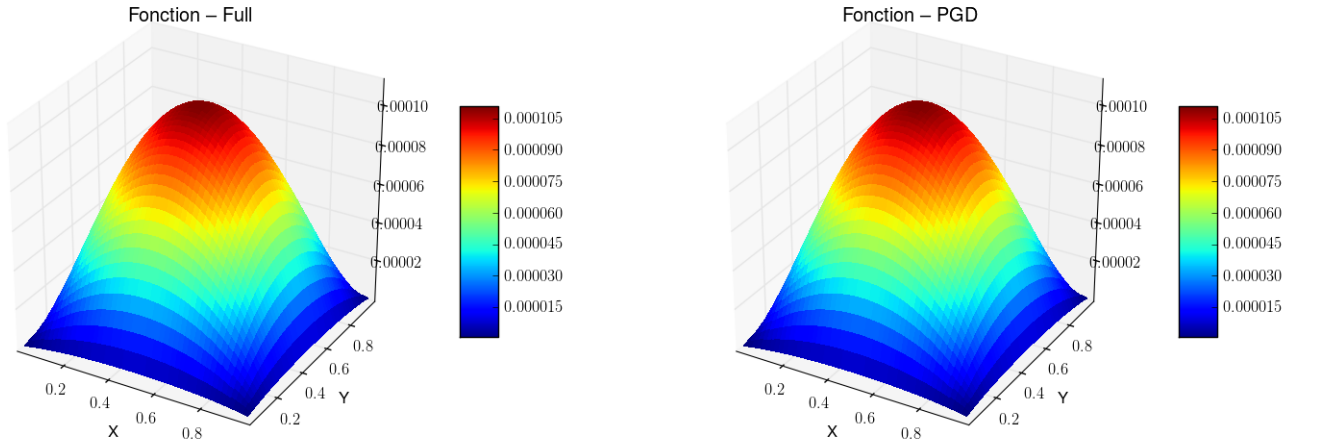


FIGURE 4.6 – Solutions obtenues par résolution complète et par résolution PGD (solutions obtenues pour  $\underline{t}_{min}$  c'est à dire là où la fonction  $F$  est convexe) – représentation 3D

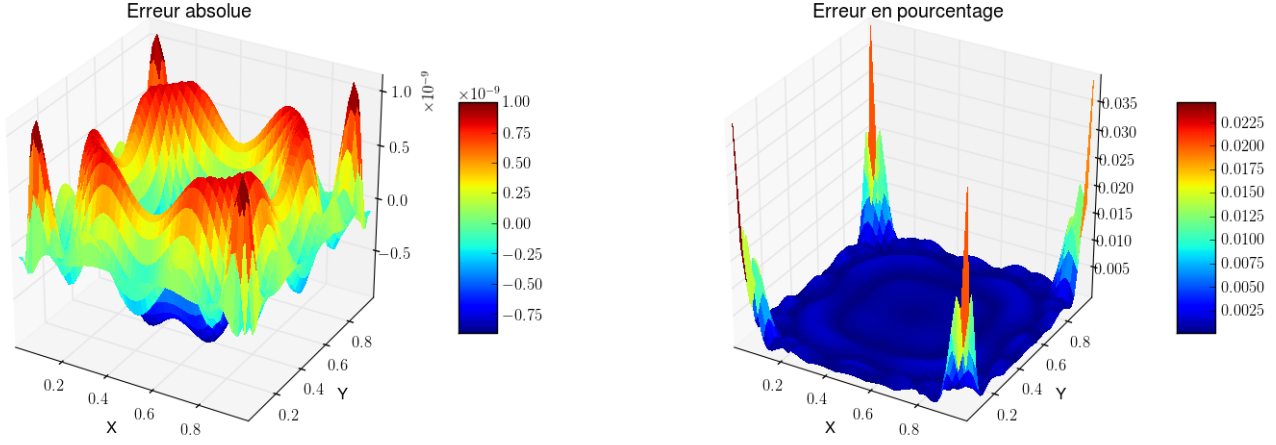


FIGURE 4.7 – Comparaison des solutions obtenues par résolution complète et par résolution PGD (solution obtenue pour  $\underline{t}_{min}$  c'est à dire là où la fonction  $F$  est convexe)– représentation 3D

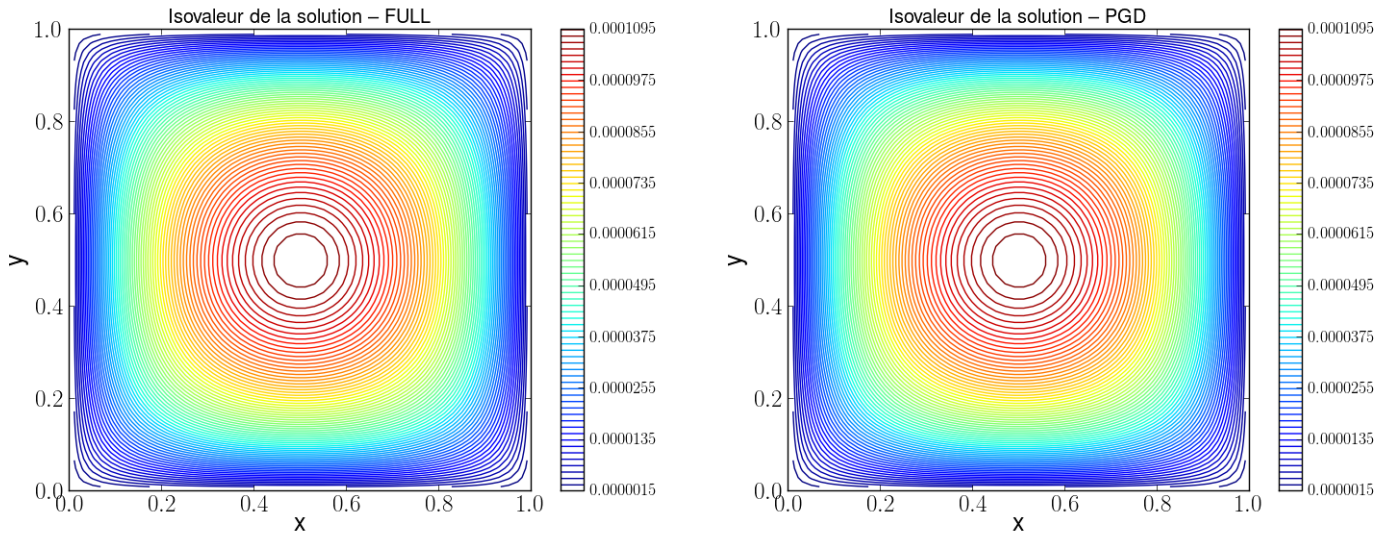


FIGURE 4.8 – Solutions obtenues par résolution complète et par résolution PGD (solutions obtenues pour  $\underline{t}_{min}$  c'est à dire là où la fonction  $F$  est convexe) – représentation 2D



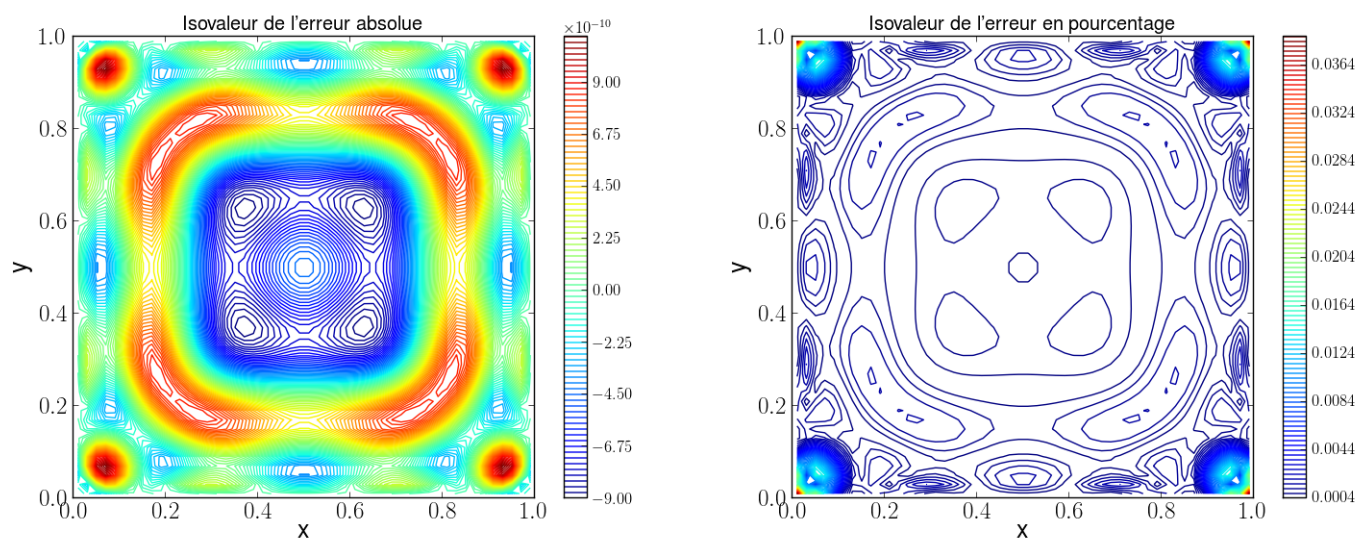


FIGURE 4.9 – Comparaison des solutions obtenues par résolution complète et par résolution PGD (solution obtenue pour  $\bar{t}_{max}$  c'est à dire là où la fonction  $F$  est convexe) – représentation 2D

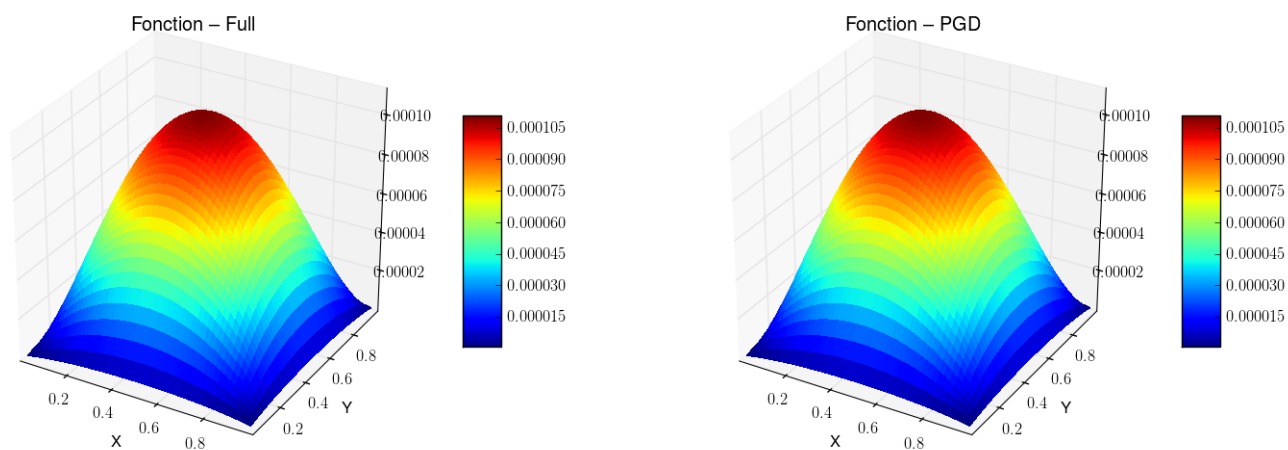


FIGURE 4.10 – Solutions obtenues par résolution complète et par résolution PGD (par méthode traditionnelle, dans ce cas on obtient uniquement la solution associée à  $t_{min}$ ) – représentation 3D

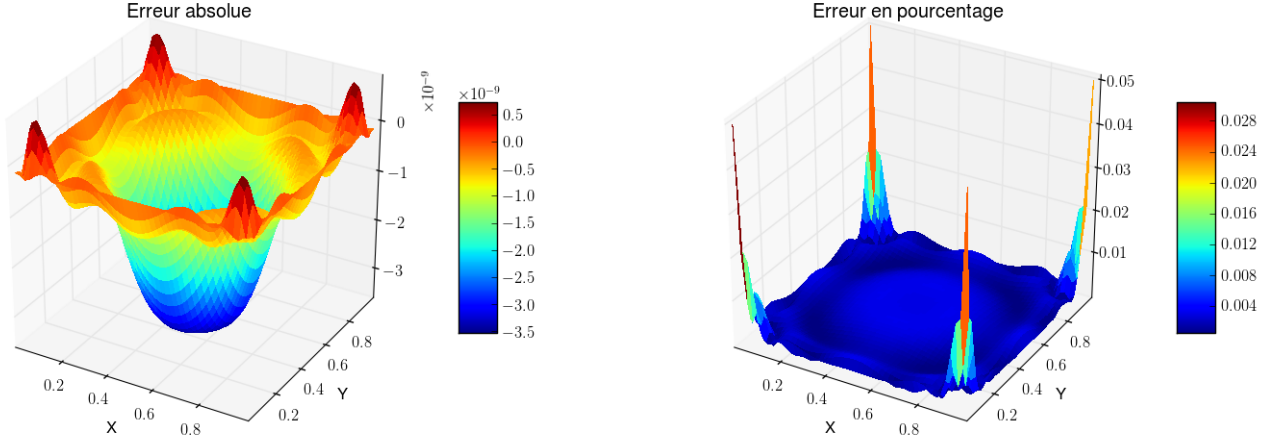


FIGURE 4.11 – Comparaison des solutions obtenues par résolution complète et par résolution PGD (par méthode traditionnelle, dans ce cas on obtient uniquement la solution associée à  $t_{min}$ ) – représentation 3D

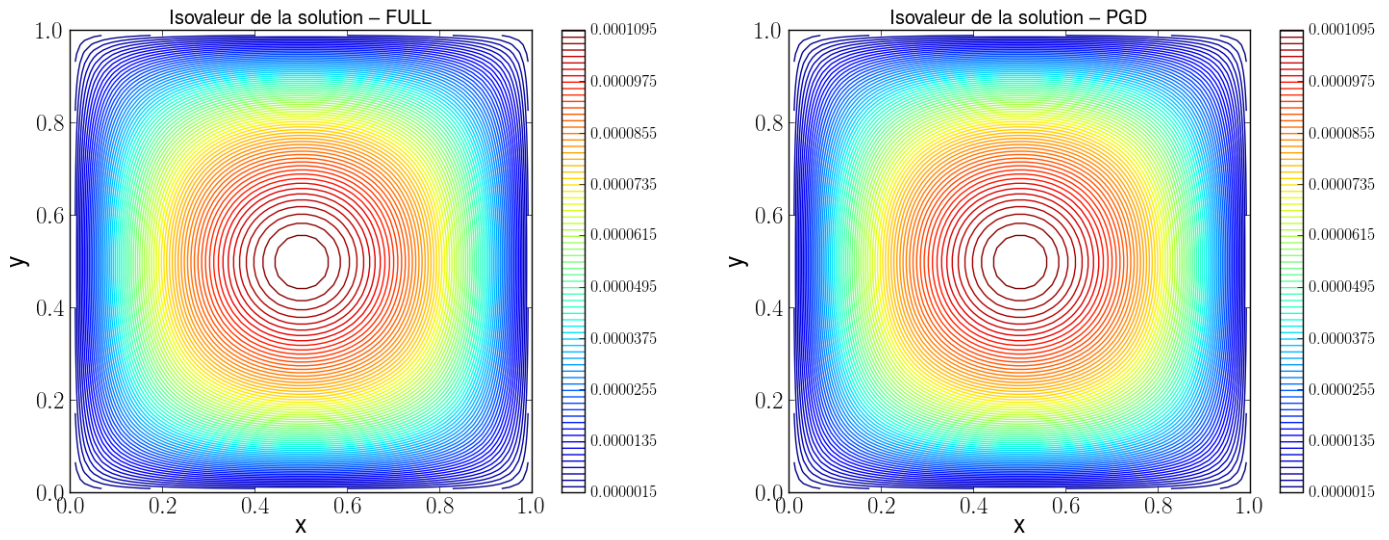


FIGURE 4.12 – Solutions obtenues par résolution complète et par résolution PGD (par méthode traditionnelle, dans ce cas on obtient uniquement la solution associée à  $t_{min}$ ) – représentation 2D

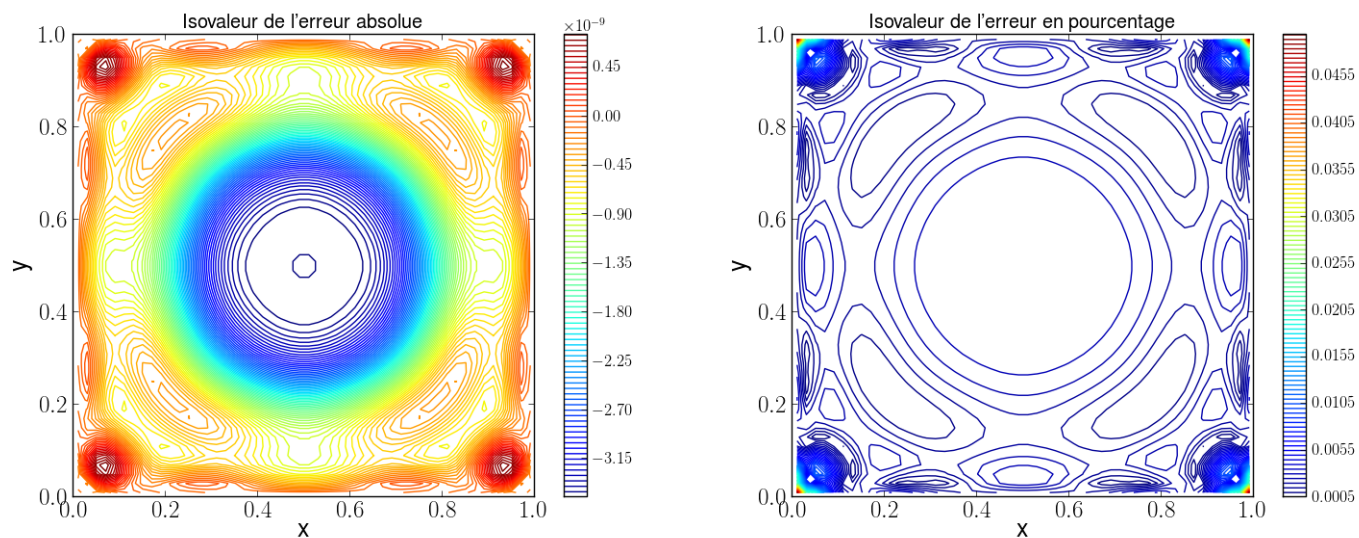


FIGURE 4.13 – Comparaison des solutions obtenues par résolution complète et par résolution PGD (par méthode traditionnelle, dans ce cas on obtient uniquement la solution associée à  $t_{min}$ )– représentation 2D

- 5) Si on fait varier le nombre de noeuds : on constate bien que plus le nombre de noeuds augmente plus la méthode PGD est rapide en comparaison avec la méthode de résolution complète.

\* — \* — \* — \* — \* — \* — \*





# Chapitre 5

## Annexe

### Sommaire

---

<b>5.1</b>	<b>Annexe</b>	<b>90</b>
5.1.1	Condition de Palais-Smale	90
5.1.2	Principe Variationnel d'Ekeland	90
5.1.3	Lemme de Brézis-Lieb	91
5.1.4	Flot Optique	91

---

## 5.1 Annexe

Dans cette partie, on va rappeler quelques définitions, théorèmes et outils essentiels que nous utilisons au cours de cette thèse.

### 5.1.1 Condition de Palais-Smale

Pour exprimer la compacité de certaines suites minimisantes, on a souvent recours à la condition de Palais-Smale (en abrégé (PS)).

**Définition 5.1.1.** Soit  $W$  un espace de Banach et  $J : W \longrightarrow \mathbb{R}$  une fonction de classe  $\mathcal{C}^1$ . On dit que la suite  $(u_n) \subset W$  est une suite de Palais-Smale s'il existe  $c \in \mathbb{R}$  tel que :

$$J(u_n) \longrightarrow c \text{ dans } \mathbb{R} \text{ et } J'(u_n) \longrightarrow 0 \text{ dans } W'.$$

On dit que  $J$  vérifie la condition de Palais-Smale au niveau  $c \in \mathbb{R}$  si toute suite  $(u_n)_n$  de Palais-Smale pour la fonctionnelle  $J$  (de niveau  $c$ ) contient une sous-suite  $(u_{n_k})_k$  convergente.

On dispose généralement un théorème important permettant la construction de suites de Palais-Smale : le principe variationnel d'Ekeland.

### 5.1.2 Principe Variationnel d'Ekeland

**Théorème 5.1.2.** Soit  $(M, d)$  un espace métrique complet et soit  $J : M \longrightarrow \mathbb{R} \cup +\infty$  semi-continue inférieurement, bornée inférieurement et  $\neq +\infty$ . Alors pour chaque  $\varepsilon > 0$ ,  $\delta > 0$  et tout  $u \in M$  avec :

$$J(u) \leq \inf_M J + \varepsilon,$$

il existe un élément  $v \in M$  minimisant strictement la fonctionnelle :

$$J_v(w) \equiv J(w) + \frac{\varepsilon}{\delta} d(v, w).$$

En plus on a :

$$J(v) \leq J(u) \text{ dès que } d(u, v) \leq \delta.$$

Comme corollaire du principe d'Ekeland, on a :

**Corollaire 5.1.3.** Si  $W$  est un espace de Banach et  $J \in \mathcal{C}^1(W)$  bornée inférieurement alors il existe une suite minimisante  $(v_n)$  pour  $J$  dans  $W$  telle que :

$$J(v_n) \longrightarrow \inf_W J, \quad DJ(v_n) \longrightarrow 0 \text{ dans } W' \text{ quand } n \longrightarrow +\infty.$$

### 5.1.3 Lemme de Brézis-Lieb

**Théorème 5.1.4.** (*Lemme de Brézis-Lieb*) Soit  $j : \mathbb{C} \longrightarrow \mathbb{C}$  une fonction continue telle que  $j(0) = 0$  et vérifie la condition suivante :

Pour tout  $\varepsilon > 0$  suffisamment petit il existe deux fonctions continues et positives  $\varphi_\varepsilon$  et  $\psi_\varepsilon$  telles que :

$$(P) \quad \forall a, b \in \mathbb{C}, \quad |j(a+b) - j(a)| \leq \varepsilon \varphi_\varepsilon(a) + \psi_\varepsilon(b).$$

Supposons que  $j$  satisfait l'hypothèse du dessus et soit  $f_n = f + g_n$  une suite de fonctions mesurables de l'ouvert  $\Omega \subset \mathbb{R}^N$  vers  $\mathbb{C}$  telles que :

- i)  $g_n \rightarrow 0$  presque partout,
- ii)  $j(f) \in L^1$ ,
- iii) il existe une constante  $C < +\infty$  indépendante de  $\varepsilon$  et  $n$  telle que  $\int \varphi_\varepsilon(g_n(x)) d\mu(x) \leq C$ ,
- iv)  $\int \psi_\varepsilon(f(x)) d\mu(x) < +\infty$  pour tout  $\varepsilon > 0$ , où  $\mu$  est la mesure de Lebesgues dans  $\mathbb{R}^N$ ,

Alors :

$$\int |j(f + g_n) - j(g_n) - j(f)| d\mu \rightarrow 0 \quad \text{quand } n \rightarrow +\infty.$$

Une application directe du lemme de Brézis-Lieb pour  $j(x) = |x|^p$  dans  $L^p(\Omega)$ , où  $1 \leq p < +\infty$ . Dans cette situation on a :

$$\int (|f_n|^p - |f_n - f|^p - |f|^p) d\mu \rightarrow 0.$$

### 5.1.4 Flot Optique

L'estimation du flot optique est une étape qui consiste au calcul du mouvement apparent d'un objet. Pour calculer le flot optique on suppose que l'intensité lumineuse d'un point reste constante sur le long de son trajet.

$$\ast - \ast - \ast - \ast - \ast - \ast - \ast$$



# Perspectives

Parmi les perspectives de recherche que nous définissons, on peut citer

1. Montrer que les suites PGD sont bien définies pour les problèmes de type concave-convexe. Ainsi, trouver des résultats de convergence pour la topologie forte de telles suites. Il s'agit d'un problème difficile, du fait que la variété de Nehari associée au problème n'est pas faiblement fermée.
2. En flot optique, considérer des énergies à croissance linéaires à l'infini, parce qu'elles sont plus performantes que les termes quadratiques et ces énergies apparaissent naturellement dans les problèmes de surfaces minimales.
3. Etudier mathématiquement et numériquement les méthodes PGD pour les EDP fractionnaires.

\* — \* — \* — \* — \* — \* — \*



# Liste des travaux

## Articles publiés

- On eigenelements sensitivity for compact self-adjoint operators and applications with A. EL HAMIDI and Aziz HAMDOUNI [Chapitre 2].
- A Progressive Generalized Decomposition approach of variational methods : Example on L2 - L2 Optical Flow with A. EL HAMIDI and Baudouin Denis de Senneville. [Chapitre 3].

## Travail en cours

- PGD in the concave-convex problem with A. EL HAMIDI and C.ALLERY. [Chapitre 4].

\* - \* - \* - \* - \* - \* - \*





# Bibliographie

- [1] N. Akkari, A. Hamdouni, E. Liberge, M. Jazar : *A mathematical and numerical study of the sensitivity of a reduced order model by POD (ROM-POD), for a 2D incompressible fluid flow*, Journal of Computational and Applied Mathematics, 270 (2014) 522-530
- [2] N. Akkari, A. Hamdouni, M. Jazar, *Mathematical and numerical results on the sensitivity of the POD approximation relative to the Burgers equation*, Applied Mathematics and Computation, 247 (2014) 951-961
- [3] C. Allery, C. Béghein, A. Hamdouni, *On investigation of particle dispersion by a POD approach*, Int. Applied Mechanics, 44 (2008) 110-119.
- [4] O. Allix, P. Vidal, *A new multi-solution approach suitable for structural identification problems*. Comput Methods Appl Mech Eng (2002) 191 :2727-2758.
- [5] C. O. Alves, *Existence of Positive Solutions for a Problem with Lack of Compactness Involving the  $p$ -Laplacian*, Nonlinear Analysis - TMA, 51 (2002), 1187–1206.
- [6] C. O. Alves, P. C. Carriao, & O. H. Miyagaki, *Signed solution for a class of quasilinear elliptic problem with critical growth*, Commun. Pure Appl. Anal. 1 (2002), no. 4, 531–545.
- [7] A. Ambrosetti, H. Brezis, G. Cerami, *Combined effects of concave and convex nonlinearities in some elliptic problems*, J. Funct. Anal. 122 (1994), no. 2, 519–543.
- [8] A. Ambrosetti, J. Garcia Azorero, I. Peral, *Multiplicity results for some nonlinear elliptic equations*, J. Funct. Anal. 137 (1996) no. 1, 219–242.
- [9] A. Ambrosetti, J. Garcia Azorero, I. Peral, *Existence and multiplicity results for some nonlinear elliptic equations : a survey*, Rend. Mat. Appl. (7) 20 (2000), 167–198.
- [10] A. Ammar, B. Mokdad, F. Chinesta, R. Keunings, *A new family of solvers for some classes of multidimensional partial differential equations encountered in kinetic theory modeling of complex fluids*, Journal of Non-Newtonian Fluid Mechanics 139 (2006) 153–176.
- [11] A. Ammar, B. Mokdad, F. Chinesta, R. Keunings, *A new family of solvers for some classes of multidimensional partial differential equations encountered in*

- kinetic theory modeling of complex fluids. Part II : transient simulation using space-time separated representations*, J. Non-Newtonian Fluid Mechanics 144 (2007) 98–121.
- [12] A. Ammar, M. Normandin, F. Daim, D. Gonzalez, E. Cueto, F. Chinesta, *Non incremental strategies based on separated representations : Applications in computational rheology*. Communications in Mathematical Sciences, 8(3, Sp. Iss. SI) (2010) :671-695.
- [13] C.A. Antoulas, *Approximation of Large-scale Dynamical Systems*. Advances in design and control. SIAM 2005.
- [14] B. Denis de Senneville, A. El Hamidi, C. Moonen, *A direct PCA-based approach for real-time description of physiological organ deformations*, Transactions on Medical Imaging, Vol. 34, No. 4, April 2015.
- [15] T. Bartsch, M. Willem, *On an elliptic equation with concave and convex nonlinearities*, Proc. Amer. Math. Soc. 123 (1995), no. 11, 3555–3561.
- [16] L. Boccardo, M. Escobedo, I. Peral, *A Dirichlet problem involving critical exponents*, Nonlinear Anal. 24 (1995), 1639–1648.
- [17] J. F. Bonder & J. D. Rossi, *Existence results for the  $p$ -Laplacian with non-linear boundary conditions*, J. Math. Anal. Appl. 263 (2001) 195–223.
- [18] J. F. Bonder & J. D. Rossi, *Asymptotic behavior of the best Sobolev trace constant in expanding and contracting domains*, Commun. Pure Appl. Anal. 1 (2002), no. 3, 359–378.
- [19] H. Brézis, E. Lieb, *A Relation Between Pointwise Convergence of Functions and Convergence of Functionals*, Proc. Amer. Math. Soc. 88 (1983) 486–490.
- [20] C. Blanzé, R. Danwe, P. Ladevèze, J-P Moreau, U *Une méthode pour l'étude d'assemblage de structures massives*. In : Colloque National en Calcul des Structures, (1993) Hermès, pp 913-919.
- [21] C. Blanzé, L. Champaney, J-Y .Cognard, P. Ladevèze, *A modular approach to structure assembly computations application to contact problems*. Eng Comput (1996) 13(1) :15.
- [22] R. Bhatia, L. Elsner, *The Hoffman-Wielandt inequality in infinite dimensions*, Proc. Indian Acad. Sci. (Math. Sci.), 104 (1994) 483-494
- [23] PA. Boucard, P. Ladevèze, M. Poss, P. Rougée, *A nonincremental approach for large displacement problems*. Comput Struct (1997) 64 :499-508.
- [24] PA. Boucard, P. Ladevèze, *A multiple solution method for non-linear structural mechanics*. Mech Eng (1999) 50(5) :317-328.
- [25] PA. Boucard, P. Ladevèze, *Une application de la méthode latin au calcul multirésolution de structures non linéaires*. In : Revue Européenne des Eléments Finis, (1999) pp 903-920

- 
- [26] P. A. Boucard, *Application of the LATIN method to the calculation of response surfaces*. In : 1st MIT conference on computational fluid and solid mechanics, (2001) vol 1, pp 78-81.
  - [27] W. Cazemier, R.W. C. P. Verstappen, et A. E. P. Veldman, *Proper orthogonal decomposition and low-dimensional models for driven cavity flows*. J. of Physics Fluids, 10 :1685-1699, 1998.
  - [28] L. Champaney, J-Y. Cognard, D. Dureisseix, P. Ladevèze, *Large scale applications on parallel computers of a mixed domain decomposition method*. (1997) Comput Mech 19(4) :253-263.
  - [29] F. Chinesta, A. Ammar, E. Cueto, *Recent Advances and New Challenges in the Use of the Proper Generalized Decomposition for Solving Multidimensional Models*. Archives of Computational Methods in Engineering 17 (2010) 327-350.
  - [30] F. Chinesta, P. Ladevèze, E. Cueto, *A Short Review on Model Order Reduction Based on Proper Generalized Decomposition*. Archives of Computational Methods in Engineering 18 (2011) 395-404.
  - [31] F. Chinesta, A. Ammar, A. Leygue, R. Keunings, *An overview of the proper generalized decomposition with applications in computational rheology*. Journal of Non-Newtonian Fluid Mechanics 166 (2011) 578-592.
  - [32] F. Chinesta, F. Ammar, F. Lemarchand, P. Beauchene, F. Boust, *Alleviating mesh constraints : Model reduction, parallel time integration and high resolution homogenization*. Computer methods in applied mechanics and engineering, 197 :400-413.
  - [33] C. Davis, W. M. Kahan, *The rotation of eigenvectors by a perturbation. III*, SIAM J. Numer. Anal., 7 (1970) 1-46.
  - [34] A. Deane, I. Kevrekidis, G. Karniadakis, S. Orszag, *Low-dimensional models for complex geometry flows : Application to grooved channels and circular cylinders*. J. of Physics Fluids, 3 :2337-2354, 1991.
  - [35] K. I. Diamataras, S. Y. Kung, *Principal Component Neural Networks : theory and Applications*. New York : John Wiley and Sons Inc.
  - [36] A. Dumon, C. Allery, A. Ammar, *Proper general decomposition (PGD) for the resolution of Navier-Stokes equations*. Journal of Computational Physics 230 (2011) 1387-1407.
  - [37] D. Dureisseix, P. Ladevèze, D. Néron, BA. Schrefler, *A multi-time-scale strategy for multiphysics problems : application to poroelasticity*. Int J Multiscale Comput Eng (2003), 1(4) :387-400.
  - [38] P. Drabek, S. Pohozaev, *Positive solutions for the  $p$ -Laplacian : application of the fibering method*, Proc. Roy. Soc. Edinburgh Sect. A 127 (1997) 703–726.

- [39] N. Akkari, A. Hamdouni, E. Liberge, M. Jazzar, *On the sensitivity of the POD technique for a parameterized quasi-nonlinear parabolic equation*, Advanced Modeling and Simulation in Engineering Sciences, 2 (2014) 1-16
- [40] E. Liberge, A. Hamdouni, *Reduced order modelling method via proper orthogonal decomposition (POD) for flow around an oscillating cylinder*, Journal of Fluids and Structures, 26 (2010) 292-311
- [41] R. Everson, L. Sirovich, *Karhunen-Loeve procedure for gappy data*. Journal of the Optical Society of America A : Optics, Image Science, and Vision, Volume 12, Issue 8, August 1995, pp.1657-1664.
- [42] T. Lieu, C. Farhat, M. Lesoinne, *Reduced-order fluid/structure modeling of a complete aircraft configuration*. Computer Methods in Applied Mechanics and Engineering 195, 41-43 (2006), 5730-5742.
- [43] J. Favier, A. Kourta, *Study of separation control on a wing profile using piv measurements and pod analysis*. Comptes Rendus - Mecanique 334, 4 (2006), 272-278.
- [44] C. Ghnatios, F. Chinesta, E. Cueto, A. Leygue, P. Breitenkopf, P. Villon, *Methodological approach to efficient modeling and optimization of thermal processes taking place in a die : application to pultrusion*. Composites Part A : Applied Science and Manufacturing 42 (2011) 1169-1178.
- [45] D. Gonzalez, A. Ammar, F. Chinesta, E. Cueto, *Recent advances on the use of separated representations*. International Journal for Numerical Methods in Engineering 81 (2010) 637-659.
- [46] A. Hay, J. Borggaard, D. Pelletier, *Improved low-order modeling from sensitivity analysis of the proper orthogonal decomposition*, J. Fluid Mech., 629 (2009) 41-72.
- [47] B.K. Horn, B.G. Schunk, *Determining Optical Flow*, Artificial Intelligence, 17 (1981), 185-203.
- [48] J. Hoffman, H. W. Wielandt, *The variation of the spectrum of a normal matrix*, Duke Math. J., 20 (1953) 37-39
- [49] D. Hömberg, S. Volkwein, *Control of laser surface hardening by a reduced-order approach utilizing proper orthogonal decomposition*, Math. Comput. Model., 38 (2003) 1003-1028.
- [50] I. T. Jolliffe, *Principal Component Analysis*, Springer, New York, 1986.
- [51] T. Kato, *Perturbation theory for linear operators*, Springer-Verlag, 1980.
- [52] K. Karhunen, *Über Lineare Methoden in der Wahrscheinlichkeitsrechnung*, Annals of Academic Science Fennicae, Series A1 Mathematics and Physics 37, 1946, 3-79.
- [53] D. Kosambi, *Statistics in function space*, Journal of Indian Mathematical Society 7, 1943, 76-88.

- 
- [54] P. Korman, *On uniqueness of positive solutions for a class of semilinear equations*, Discrete Contin. Dyn. Syst. 8 (2002), no. 4, 865–871.
  - [55] K. Kunisch, S. Volkwein, *Galerkin proper orthogonal decomposition methods for a general equation in fluid dynamics*, SIAM J. Numer. Anal., 40 (2002) 492-515
  - [56] K. Kunisch, S. Volkwein, *Control of Burgers equation by a reduced order approach using proper orthogonal decomposition*, J. Optim. Theory Appl., 102 (1999) 345-371
  - [57] P. Ladevèze, Ph. Lorong, *A large time increment approach with domain decomposition technique for mechanical non linear problems*. In : Computing methods in applied sciences and engineering INRIA, (1992) pp. 569-578.
  - [58] P. Ladevèze, Ph. Lorong, *Formulation et stratégies "parallèles" pour l'analyse non linéaire des structures*. In : Colloque national en calcul des structures. (1993) Hermès, Paris, pp 910-919.
  - [59] P. Ladevèze, *Mécanique non linéaire des structures*. (1996) Hermès, Paris.
  - [60] P. Ladevèze, D. Dureisseix, *A 2-level and mixed domain decomposition approach for structural analysis*. Contemp Math (1998) 218 :246-253.
  - [61] P. Ladevèze, D. Dureisseix, *A new micro-macro computational strategy for structural analysis*. CR Acad Sci, Ser II, Fascicule, B-Mec Phys Astron, (1999) 327(12) :1237-1244.
  - [62] P. Ladevèze. *Nonlinear Computational Structural Mechanics*. Springer, N.Y., 1999.
  - [63] P. Ladevèze, *New algorithms : mechanical framework and development*. Technical Report 57, (1985) LMT-Cachan.
  - [64] P. Ladevèze, *On a family of algorithms for structural mechanics*. CR Acad Sci Paris (1985) 300(2) :41-44.
  - [65] P. Ladevèze, O. Loiseau, D. Dureisseix, *A micro-macro and parallel computational strategy for high heterogeneous structures*. Int J Numer Methods Eng, (2001) 52(1-2) :121-138.
  - [66] P. Ladevèze, J.C. Passieux, D. Néron, *The LATIN multiscale computational method and the Proper Generalized Decomposition*, Computer Methods in Applied Mechanics and Engineering 199 (2010) 1287-1296.
  - [67] P. Ladevèze, *Sur une famille d'algorithmes en mécanique des structures*. Compte rendu de l'académie des sciences, 300(2) :41-44.
  - [68] T. Lassila, G. Rozza, *Parametric free-form shape design with PDE models and reduced basis models*, Comput. Methods Appl. Mech. Engrg., 199 (2010) 1583-1592
  - [69] A. Leygue, E. Verron, *A first step towards the use of proper general decomposition method for structural optimization*, Archives of Computational Methods in Engineering 17 (2010) 465-472.

- [70] C. G. Wu, Y. C. Liang, W. Z. Lin, H. P. Lee, S. P. Lim, *A note on equivalence of proper orthogonal decomposition methods*, Journal of Sound and Vibration 265, 2003, 1103-1110.
- [71] M. Loeve, *Fonctions Aléatoires du Second Ordre*, in Processus stochastiques et mouvement Brownien, P. Levy (ed.), Gauthier-Villars, Paris, 1948.
- [72] C. Long, *Visualisation of matrix singular value decomposition*. Mathematics Magazine, (1983), 56 :161-167.
- [73] J. L. Lumley, *The structure of inhomogeneous turbulent flows*. In A.M Iaglom and V.I Tatarski, editors, Atmospheric Turbulence and Radio Wave Propagation, pages 221-227. Nauka, Moscow, 1967.
- [74] D. Mumford, J. Shah, *Optimal approximations by piecewise smooth functions and associated variational problems*, Comm. Pure Appl. Math., 42 (1989) 577-684.
- [75] Z. Nehari, *On a class of nonlinear second-order differential equations*, Trans. Amer. Math. Soc. 95 (1960) 101-123.
- [76] D. Néron, D. Dureisseix, *A computational strategy for poroelastic problems with a time interface between coupled physics*. Int J Numer Methods Eng (2008), 73(6) :783-804.
- [77] D. Néron, D. Dureisseix, *A computational strategy for thermo-poroelastic structures with a time-space interface coupling*. Int J Numer Methods Eng (2008), 75(9) :1053-1084.
- [78] S. Niroomandi, I. Alfaro, E. Cueto, F. Chinesta, Model order reduction for hyperelastic materials, International Journal for Numerical Methods in Engineering 81 (2010) 1180-1206.
- [79] A. Nouy. *A generalized spectral decomposition technique to solve a class of linear stochastic partial differential equations*. Computer Methods in Applied Mechanics and Engineering, 196(45-48) :4521-4537, 2007.
- [80] P. Ladevèze, A. Nouy, *On a multiscale computational strategy with time and space homogenization for structural mechanics*. Compu. Methods Appl. Mech. Engrg, (2003), 192.
- [81] A. Nouy, P. Ladevèze, *Multiscale computational strategy with time and space homogenization : a radial-type approximation technique for solving micro problems*. Int. J. Multiscale Comput. Engrg, (2004), 170.
- [82] A. Nouy, *A generalized spectral decomposition technique to solve a class of linear stochastic partial differential equations*. Comput.Meth.Appl.Mech.Engrg, (2007), 196 :4251-4537.
- [83] A. Nouy, *Generalized spectral decomposition method for solving stochastic finite element equations : invariant subspace problem and dedicated algorithms*. Comput.Meth.Appl.Mech.Engrg, (2008), 197 :4718-4736.

- 
- [84] A. Nouy, O. Le Maître, *Generalized spectral decompositions for stochastic nonlinear problems*. J.Comput.Phys, (2007), 8 :283-288.
  - [85] M. A. Obukhov, *Statistical description of continuous fields*, Transactions of the Geophysical International Academy Nauk USSR 24, 1954, 3-42.
  - [86] M. Pomarède, *Investigation et application des méthodes d'ordre réduit pour les calculs d'écoulements dans les faisceaux tubulaires d'échangeurs de chaleur*, PhD thesis, University of La Rochelle (2012).
  - [87] M. Rajaei, S. Karlson, L. Sirovich, *Low-dimensional description of free shear flow coherent structures and their dynamical behavior*. J. of Fluid Mechanics, 258 :1401-1402, 1994.
  - [88] S.S. Ravindran *A Reduced Order Approach to Optimal Control of Fluids using Proper Orthogonal Decomposition*, International Journal for Numerical Methods in Fluids , Vol. 34 (5), pp. 425-448, 2000.
  - [89] D. Rempfer, *Investigations of boundary layer transition via galerkin projection of empirical eigenfunctions*. J. of Physics Fluids, 8 :175-188, 1996.
  - [90] L. Rudin, S. Osher, E. Fatemi, *Nonlinear total variation based noise removal algorithms*, Phys. D, 60 (1992) 259-268.
  - [91] S. Roujol, M. Ries, B. Quesson, C. Moonen, B. Denis de Senneville, *Real-time MR-thermometry and dosimetry for interventional guidance on abdominal organs*, Magnetic Resonance in Medicine, (2010).
  - [92] B. Rousselet, D. Chenais, *Continuité et différentiabilité d'éléments propres : Application à l'optimisation de structures*, Appl. Math. Optim., 22 (1990) 27-59
  - [93] L. Sirovitch, *Turbulence and the dynamics of coherent structures, part i : Coherent structures, part ii : Symmetries and transformations, part iii : Dynamics and scaling*. Quarterly of Applied Mathematics 45(3) (1987), 561-590.
  - [94] S. Volkwein, *Optimal control of a phase-field model using the proper orthogonal decomposition*, Z. Angew. Math. Mech., 81 (2001) 83-97
  - [95] S. Watanabe, *Karhunen-Loève expansion and factor analysis theoretical remarks and applications*, in Proceedings of the 4th Conference on Information Theory, Prague, Czech Republic, 1965.
  - [96] M. Willem, *Minimax theorems*, Progress in Nonlinear Differential Equations and their Applications, 24. Birkhauser Boston, Inc., Boston, MA, (1996).
  - [97] Y. C. Liang, W. Z. Lin, C. G. Wu, *Proper orthogonal decomposition and its applications*, Journal of Sound and Vibration 2002, 527-544.







---

*Mathematical and numerical study of POD and PGD dimensional reduction methods*

---

**Abstract** This thesis is formed of four chapters. The first one presents the mathematical notions and tools used in this thesis and gives a description of the main results obtained within.

The second chapter presents our generalization of a result obtained by Rousselet-Chenais in 1990 which describes the sensitivity of eigensubspaces for self-adjoint compact operators. Rousselet-Chenais were limited to sensitivity for specific subspaces of dimension 1, we have extended their result to higher dimensions. We applied our results to the *Proper Orthogonal Decomposition* (POD) in the case of parametric, temporal and spatial variations (Gappy-POD).

The third chapter discusses the optical flow estimate with quadratic or linear energies at infinity. Mathematical results of convergence are shown for the method *Progressive Generalized Decomposition* (PGD) in the case of quadratic energies. Our proof is based on the decomposition of Brézis-lieb via the convergence almost everywhere of the PGD sequence gradients. A detailed numerical study is made on different types of images : on the passive scalar transport equations, whose displacement fields are solutions of the Navier-Stokes equations. These equations present a challenge for optical flow estimates because of the presence of low gradient regions in the image. We applied our method to the MRI image sequences to estimate the movement of the abdominal organs. PGD presented a superiority in both computing time level (even in 2D) and accuracy representation of the estimated motion. The local diffusion of standard methods (Horn Schunck, for example) limits the convergence rate, in contrast to the PGD which is a more global approach by construction. The last chapter deals with the application of PGD method in the case of variational elliptic equations whose energy present all challenges to classical variational methods : lack of convexity, lack of coercivity and lack of boundedness. We prove convergence results for the weak topology, the PGD sequences converge (when they are well defined) to two extremal solutions on the Nehari manifold. Several mathematical questions about PGD remain open in this chapter. These questions are part of our research perspectives.

---

**Keywords :** Proper Orthogonal Decomposition (POD), parametric sensitivity, Proper Generalized Decomposition (PGD), Reduced order models (ROM), Nehari manifold, Optical Flow.

---

---

*Etude mathématique et numérique des méthodes de réduction dimensionnelle de type POD et PGD*

---

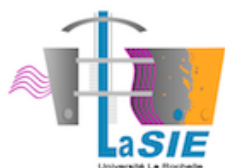
**Résumé** Ce mémoire de thèse est formé de quatre chapitres. Un premier chapitre présente les différentes notions et outils mathématiques utilisés dans le corps de la thèse ainsi qu'une description des résultats principaux que nous avons obtenus.

Le second chapitre présente une généralisation d'un résultat obtenu par Rousselet-Chénais en 1990 qui décrit la sensibilité des sous-espaces propres d'opérateurs compacts auto-adjoints. Rousselet-Chénais se sont limités aux sous-espaces propres de dimension 1 et nous avons étendu leur résultat aux dimensions supérieures. Nous avons appliqué nos résultats à la *Décomposition par Projection Orthogonale* (POD) dans le cas de variation paramétrique, temporelle ou spatiale (Gappy-POD). Le troisième chapitre traite de l'estimation du flot optique avec des énergies quadratiques ou linéaires à l'infini. On montre des résultats mathématiques de convergence de la méthode de *Décomposition Progressive Généralisée* (PGD) dans le cas des énergies quadratiques. Notre démonstration est basée sur la décomposition de Brézis-Lieb via la convergence presque-partout de la suite gradient PGD. Une étude numérique détaillée est faite sur différents type d'images : sur les équations de transport de scalaire passif, dont le champs de déplacement est solution des équations de Navier-Stokes. Ces équations présentent un défi pour l'estimation du flot optique à cause du faible gradient dans plusieurs régions de l'image. Nous avons appliqué notre méthode aux séquences d'images IRM pour l'estimation du mouvement des organes abdominaux. La PGD a présenté une supériorité à la fois au niveau du temps de calcul (même en 2D) et au niveau de la représentation correcte des mouvements estimés. La diffusion locale des méthodes classiques (Horn & Schunck, par exemple) ralentit leur convergence contrairement à la PGD qui est une méthode plus globale par nature. Le dernier chapitre traite de l'application de la méthode PGD dans le cas d'équations elliptiques variationnelles dont l'énergie présente tous les défis aux méthodes variationnelles classiques : manque de convexité, manque de coercivité et manque du caractère borné de l'énergie. Nous démontrons des résultats de convergence, pour la topologie faible, des suites PGD (lorsqu'elles sont bien définies) vers deux solutions extrémales sur la variété de Nehari. Plusieurs questions mathématiques concernant la PGD restent ouvertes dans ce chapitre. Ces questions font parties de nos perspectives de recherche.

---

**Mot-clefs :** Décoposition par projection orthogonale (POD), sensibilité paramétrique, Décomposition Propre Généralisée (PGD), Réduction de modèles (ROM), Variété de Nehari, Flot Optique.

---



LASIE  
Laboratoire des Sciences de l'Ingénieur pour  
l'Environnement, Avenue Michel Crépeau  
17042 LA ROCHELLE